A versatile method for the coupling of protein to DNA: synthesis of α -macroglobulin-DNA conjugates

Sheue-yann Cheng, Glenn T.Merlino and Ira H.Pastan

National Cancer Institute, Laboratory of Molecular Biology, Bethesda, MD 20205, USA

Received 10 December 1982; Accepted 12 January 1983

ABSTRACT

We describe a simple, general method to link proteins covalently to DNA. The method uses two reagents, N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine and 2-iminothiolane. The former reacts specifically with nonpaired quanine residues and upon reduction generates a free sulfhydryl group. The latter reacts with a protein to provide another sulfhydryl group which is subsequently conjugated to DNA by an intermolecular disulfide interchange reaction. Using this method a2-macroglobulin was conjugated to plasmid DNA encoding the Herpes simplex virus-1 thymidine kinase gene or a DNA fragment containing the E. coli chloramphenicol acetyltransferase gene. Up to 20% of the total DNA was conjugated to a2-macroglobulin and the a2-macroglobulin-DNA conjugate had a protein/DNA molar ratio of approximately two. The whole reaction takes place under very mild conditions in aqueous solution. The structure of DNA appears not to be significantly affected by the chemical modification. This method may prove useful in ligand directed gene transfer studies.

INTRODUCTION

One of the medical applications of gene splicing technology could be the introduction of DNA into cells with genetic defects. One way to accomplish this is to direct DNA to cells within an intact organism by linking the DNA to a protein that will be bound by a specific receptor on the cell surface and subsequently be internalized. Thus by coupling a gene to a protein ligand, one could confer receptor-selectivity to that gene.

Several different approaches have been used to crosslink proteins to nucleic acids: these include ultraviolet irradiation (1-3), bifunctional chemical (4, 5) and photochemical cross-linking (6-8). Among the latter, phenylglyoxyl derivatives are widely used for protein-nucleic acid cross linking. Under mild conditions, glyoxyl derivatives react specifically with nonpaired quanine residues to form adducts with known structures (9-10). Thus, glyoxyl derivatives have been extensively used as probes to study the structures of ribosomes (11-15). Recently, Expert-Bezancon and Hayer described the synthesis of N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine (16). In the present study, we

describe the use of this compound as a bifunctional reagent to couple plasmid DNA encoding either the Herpes simplex virus-1 (HSV-1)thymidine kinase gene (LE578) or the E. coli chloramphenicol acetyltransferase gene (CAT) to a2 macroglobulin (a2M) via mixed disulfide formation.

EXPERIMENTAL PROCEDURES

Materials

Selenium dioxide, 4-acetylbenzoic acid, and acetic anhydride were purchased from Aldrich. 5,5'-dithio-bis-(2-nitrobenzoic acid) and 2-iminothiolane hydrochloride were obtained from Pierce. Dithiothreitol was from Bethesda Research Laboratories. Activated thiol-Sepharose 4B was from Pharmacia Fine Chemicals. The plasmid LE578, containing the Bam Hi DNA fragment encoding HSV-1 thymidine kinase (17) was a gift from L. Enquist (Molecular Genetics, Minneapolis, Minnesota). The plasmid pASV-cat containing the CAT gene (18) was generously provided by C. Gorman and B. Howard (NIH). Kpn I, Nde I, and Bam HI were purchased from New England Biolabs. Terminal transferase was obtained from P-L Biochemicals and $\lceil \alpha - \frac{32p}{dG} \rceil$ was from New England Nuclear (800 Ci/mmol). Buffers

(1) Buffer A: 0.05 M triethanolamine hydrochloride, 15 mM boric acid, pH 7.8

(2) Buffer B: 0.05 M triethanolamine hydrochloride, 15 mM boric acid, 50 mM NaCl, 50 mM dithiothreitol, pH 7.8

(3) Buffer C: 0.05 M triethanolamine hydrochloride, 15 mM boric acid, 0.1 M-NaCl, ¹ mM EDTA, pH 7.8

(4) Buffer D: 0.05 M triethanolamine hydrochloride, 15 mM boric acid, 0.1 M NaCl, ¹ mM EDTA, 50 mM dithiothreitol, pH 8.2

(5) Buffer E: 0.014 M NaCl, 50 mM phosphate, ¹ mM EDTA, 0.1% bovine serum albumin, pH 7.4

Synthesis of N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine

The intermidiate compound, p-carboxyphenylglyoxal was synthesized by the oxidation of p-acetylbenzoic acid with selenium dioxide according to (19, 20). p-Carboxyphenylglyoxal was further reacted with cystamine and acetic anhydride to form N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine according to (16). The final purification step was modified as follows. After the reaction as described in (16), the reaction product (yellow oily residue) was applied to a silica gel column (3 cm \times 35 cm) using 1-butanol: acetic acid:water (5:3:2 v/v) as eluant. The eluate (flow rate: 0.5 ml/min; 2 ml/fraction) was monitored by the absorption at 350 nm. Fractions 88-91 were pooled and evaporated to dryness. The dry residue was further purified by preparative thin-layer chromatography (Whatman, PLK5F) using a solvent system of benzene: acetic acid: water = $24:2:4$ (v/v/v). The band with a Rf of 0.38 was scraped and the pure product was eluted with 30% methanol in ethylacetate. The purity of this compound was confirmed by thin-layer chromatography (Whatman, MK6F) in two solvent systems: (A) benzene: acetic acid:methanol = $24:2:4$ (v/v/v); (B) ethylacetate:acetic acid = $9.5:0.5$ (v/v) with a Rf of 0.35 and 0.21, respectively. This compound has an Σ_{25}^{Lcm} = 2,940 and $\Sigma_{552}^{\text{1cm}}$ = 4,552 in aqueous and absolute methanol solution, respectively. The structure was further confirmed by PMR in CD30D:

0 $-C-CH_3$ (2.00 ppm); $-CH_2CH_2-S-CH_2CH_2-$ (3.0 ppm); phenyl protons (8.2 ppm, 0 quartet); HC-C- (8.1 ppm). I 0

Preparation and Derivatization of $[$ ¹²⁵I] α ₂-Macroglobulin

 α ₂M was prepared from whole human plasma as described previously (21, 22). a2M was iodinated using the chloramine-T procedure (23). The specific activity of $\left[\frac{125}{\text{kg}}\right]$ a₂M was 1 µCi/µg.

 α ₂M was derivatized according to (24). $[1251] \alpha$ ₂M (5.8 mg; 1.5 µCi/mg) in 1.2 ml of 10 mM Tris was dialyzed against 50 mM phosphate buffer, pH 7.6 for 20 hrs at 0°C. After filtering through a 0.45 µM Millipore Unit (Swinnex-HA), the $[1251]$ a₂M solution was cooled to 0°C and 2 mg of 2-iminothiolane hydrochloride in 0.14 ml H₂O was added. After incubation for 10 min at 0° C, the excess iminothiolane was separated from protein by a Pharmacia prepacked column PD-10. Two fractions containing the highest concentration of α 2M were pooled (2.7 mg in 1.1 ml) and 60 pl of 0.2 M 5,5'-dithio-bis-(2-nitrobenzoic acid) in 2 M phosphate, pH 7.6 was added to give a final concentration of 10 mM. After standing for 30 min at 22°C, the degree of derivatization of α 2M was quantified by measuring the absorbance at 412 nm ($\Sigma_{412}^{1cm} = 13,600$ M⁻¹) (25). The excess 5,5'-dithio-bis-(2-nitrobenzoic acid) was separated from the protein on ^a PD-10 column as described above. Two radioactive peak-fractions were pooled (1.8 mg in 0.9 ml) and used for subsequent conjugation of α ₂M to G-tailed DNA. Addition of Homopolymer Tracts to Linear DNA

The 8.1 kb plasmid LE578 (150 μ g), containing the HSV-1 thymidine kinase gene, was cleaved by approximately 250 units of Kpn ^I as described by the manufacturer. The resulting DNA was extracted once with an equal volume of phenol/ chloroform/isoamyl alcohol (24:24:1), and once with an equal volume of chloroform/isoamyl alcohol (24:1). The aqueous phase was made 0.3 M with sodium acetate (pH 5.5), precipitated with two volumes of ethanol, resuspended in water, and again restricted with Kpn ^I as above to ensure complete digestion. The addition of dGTP to the linearized plasmid LE578 was achieved using the procedure essentially as described by Roychoudhury et al. (26). The 700 µl reaction mixture contained 50 mM Tes (N-tris[hydroxymethyl]-methyl-2-aminoethane sulfonic acid) (pH 7.4), 40 mM KCl, 2.0 mM β -mercaptoethanol, 1.0 mM CoCl₂, 400 U/ml terminal transferase, 15-20 mM cold dGTP, $35-150$ μ Ci/ml $\left[\alpha-32P\right]$ dGTP, and 80 pmoles of 3' hydroxyl ends/ml LE578 plasmid DNA (approximately 220 µg/ml). The mixture was incubated for 30 min at 37°C, after which the reaction was terminated by heating to 70°C for 5 min, and extracted once with an equal volume of phenol. After centrifugation, the aqueous phase was chromatographed over G-100 Sephadex previously equilibrated with 10 mM Tris (pH 8.0), ¹ mM EDTA to remove unincorporated $[a^{-32}P]$ dGTP. The DNA in the void volume was ethanol precipitated two times, dried in vacuo, and dissolved in 60 ul of buffer A for conjugation to a2M (see below) .

A similar procedure was utilized for addition of homopolymer tracts to the DNA fragment containing the CAT gene. The CAT gene was originally obtained from the E. coli transposable element Tn9, which was cloned into a pBR322-SV40 hybrid plasmid containing the Rous sarcoma virus LTR promoter (pASV-cat, 18). The plasmid pASV-cat was restricted by Nde ^I and BamHI, and the resulting 2.2 kb DNA fragment isolated, tailed, and prepared for conjugation as described above.

Conjugation of a₂-Macroglobulin to G-tailed DNA

The reaction of N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine to unpaired quanine residues of G-tailed DNA was carried out as described in (16). To a solution of G-tailed DNA in 60 ul of buffer A was added 2.6 ul of 0.51 M N-acetyl-N'-(p-glyoxylylbenzoyl)cystamin in absolute methanol. After the solution was kept at 23°C for 2 hr, the DNA was precipitated two times by ethanol and dried in vacuo. The modified DNA was redissolved in 50 µl of buffer B and kept at 23°C for 6 min. The reduced DNA was then precipitated two times by ethanol was dried for 20 min. The precipitation step also serves the purpose of removing excess dithiothreitol. The dry reduced DNA was dissolved in 60 μ l of buffer C. An aliquot (10 μ l) was removed for assaying the degree of the conjugation (see below). The remaining DNA was precipitated by ethanol, dried and to it was added the derivatized α_2 M (see above and Fig. 1). The solution was kept at 23°C for 18 hr, and the degree of conjugation was assessed by absorption at 412 nm which which quantifies the release of 4-mercapto-2-nitrobenzoic acid. The excess $[1251]$ a2M was separated from the DNA-a2M conjugate by Sepharose 4B chromatography using buffer E as eluant.

Electrophoretic Fractionation of DNA-Protein Complexes

The labeled DNA-protein mixture was loaded directly from the column onto agarose or polyacrylamide gels. Agarose gel electrophoresis was performed in 0.7% agarose in running buffer containing 40 mM Tris-acetic acid (pH 7.9), 20 mM NaAcetate, 2 mM EDTA. The gel was dried and subjected to standard autoradiographic techniques. Alternatively, samples were electrophoresed on a 3% over 3.5% stacking polyacrylamide gel system in 40 mM Tris-HCl (pH 8.3), 1.25 mM EDTA, 44.5 mM Boric acid, the gel again dried and subjected to autoradiography. Determination of the Anount of N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine Conjugated to G-tailed DNA

Activated thio-Sepharose 4B was swollen in water at 23°C for 4 hr. A small column (0.5 \times 7 cm) of thio-Sepharose 4B was equilibrated with buffer C and an aliquot of reduced DNA (see above) was applied to the column. The column was washed with buffer C until no more radioactivity was detected ('15-20 ml). The reduced DNA was then eluted with buffer D and the radioactivity of the eluate was determined.

RESUL TS

Figure 1 shows the sequential steps in the conjugation of α ²M to G-tailed DNA. The strategy lies in the introduction of one or more sulfhydryl groups in the DNA which is then conveniently conjugated to a protein by the formation of mixed disulfide. The introduction of sulfhydryl groups into DNA was easily accomplished by reaction of the DNA with N-acetyl-N'-(p-glyoxylylbenzoyl) cystamine. The reaction was carried out in the presence of borate in order to increase the stability of the adduct formed (see Fig. 1) as shown by Expert-Begancon and Hayer (16). This reagent not only provides glyoxal groups to form adduct with the unpaired quanine residues in DNA but also generates a sulfhydryl group upon reduction. The unpaired quanine residues were easily introduced into DNA by using dGTP and terminal transferase. In the present study, ~10-25 quanine residues were added to each 3'-OH end of DNA. Because N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine only reacts with unpaired quanine residues, the coding capacity of the gene of interest is not affected. Thus, this conjugation method is quite general and should be applicable for conjugating many proteins to DNA.

In addition to the generality of the method, the reaction scheme shown in Figure ¹ provides easy assay methods at each step. The number of dGTP trans-

ferred to DNA was easily determined by $32p$ measurement. The degree of modification of the quanyl residues in DNA by N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine was conveniently monitored by using thio-Sepharose 4B. The number of sulfhydryl groups introduced into α_2M and the degree of the conjugation in the final reaction of α ₂M with DNA was easily determined by measuring the release of 4-mercapto-2-nitrobenzoic acid using absorption at 412 nm. Alternatively, the degree of conjugation of α 2M to DNA can also be estimated by measuring the 125 I/32p ratio after separation of the α ₂M-DNA conjugate from unreacted α ₂M by Sepharose 4B. As shown in Figure 2, the α ₂M-LE578 conjugate was well separated

Figure 2. Separation of the unreacted α ₂M from α ₂M-DNA conjugate by Sepharose 4B column chromatography, column size: 1.5 x 60 cm; eluate: buffer ^E (0.014 M NaCli, 50 mM phosphate, ¹ mM EDTA, 0.1% bovine serum albumin, pH 7 .4); flow rate: 8 ml/hr; fraction size: 1.4 ml/fraction.

from unreacted α_2 M. Using DNA and α_2 M ranging from 90 μ g-110 μ g and 0.9 mg-2.3 mg, respectively, the ratio of α 2M/DNA in the α 2M-DNA conjugate was 2.5 ± 0.9 (Mean \pm S.D., $n = 5$) as determined either by absorption at 412 nm or 125I/32p ratio.

DNA	Yield
	$(x \text{ of total DNA})$
G-tailed DNA $(90-110 \text{ µg})$ a	(100)
DNA reacting with N-acetyl-N'-(p-glyoxylylbenzoyl)cystamine	$60 - 70$
DNA conjugated to a2M	$16 - 20$

Table 1. Yield of derivatized DNA LE578

aRange of five different preparations.

Figure 3. Autoradiogram of G-tailed LE578 plasmid DNA before and after conjunction. The LE578 DNA was linearized by digestion with <u>Kpn</u> 1, and ³²P-labeled dGTPs were added to the ³' OH ends as described in Experimental Procedures. DNA-containing samples were extracted with phenol/chloroform, ethanol precipitated, and electrophoresed on 1% agarose. Lanes: (1) LE578 plasmid DNA tailed using $\lceil \alpha - \frac{32p}{\log T} \rceil$, (2) LE578 α ₂M conjugate pretreated with β -mercaptoethanol before phenolization to reduce disulfide bonds, thereby separating protein from DNA.

Table 1 shows a summary of the yields in the conjugation of α_2 M to LE578 DNA. The overall yield ranges from 16 to 20% of the total DNA. In the preliminary experiments, when GTP rather than dGTP was used to add to the 3'-OH end of DNA, only an overall yield of $0.1%$ DNA- α ₂M conjugate was obtained. The reason for the low yield was probably due to the degradation of the G-tails of the DNA during conjugation by ribonuclease present as a contaminant in the α ^{2M} preparation. Therefore dGTP was used in the present study.

Figure 3 compares G-tailed LE578 plasmid DNA before and after conjugation to α_2 M. It is clear that the size of the DNA was not affected by the mild reaction conditions used in the modification of DNA and its subsequent conjugation to α_2 M, suggesting that the overall structure of the LE578 DNA remains intact.

Having successfully coupled α ₂M to LE578 DNA we next determined if other

Figure 4. Autoradiogram showing the effect of β -mercaptoethanol on the α ₂M-CAT gene conjugate. The conjugate was electrophoresed after column chromatography without prior phenolization on either 0.7% agarose (A) or 3.5% polyacryl-
amide (B). Lanes: (a) 6,000 cpm of the DNA fragments from plasmid pASV-cat x amide (8). Lanes: (a) 6,000 cpm of the DNA fragments from plasmid pASV-cat ^x Nde ^I ^x Bam HI, enriched in the 2.2 kb species and tailed with [a-32P]dGTP; (b) $300,000$ cpm of 125 I-labeled, purified α 2M; (c) 500 cpm of the α 2M-CAT DNA conjugate untreated; (d) 500 cpm of the conjugate treated with 100 mM Smercaptoethanol for 30 min (37%C) prior to electrophoresis; (e) same as (a) except 1,000 cpm were loaded; (f) same as (b); (g) same as (c) except 1,100 cpm were loaded; (h) same as (d) except 1,100 cpm were loaded. The vast majority of the bands displayed in lanes c, d, g, and h result from ³²P irradiation, not 125 I. The squares (\bullet) indicate the location of the readily detectable DNA species which include the ⁵ kb partially digested pASV-cat plasmid and two DNA fragments of 2.8 and 2.2 kb. For easy comparison, lanes (c) and (d) are exposed for ^a longer time period than (a) and (b).

DNAs would lend themselves readily to this conjugation technique. The α ₂M was coupled to the 2.2 kb DNA fragment containing the E_0 . coli CAT gene using the same method with a slightly higher yield of 44% and a molar ratio α ₂M/DNA = 1.5. The fact that conjugation was achieved using ^a variety of DNA moieties indicates the versatility of the technique.

In order to definitively prove that α_2 M was conjugated to the CAT gene by disulfide linkage and not by some unknown mechanism, we electrophoresed the a2M-CAT gene conjugate on native 0.7% agarose or 3.5% polyacrylamide gels with and without prior exposure to 100 mM B-mercaptoethanol. The autoradiograms shown in Figure 4 (A and B) reveals that 8-mercaptoethanol treatment results in a reduction in the size of diffuse ^{32p}-labeled high molecular weight DNAprotein complexes (lanes c and g) to sizes equivalent to original labeled DNA fragments (lanes d and h). Further proof of the nature of the protein-DNA linkage was obtained by incubating the α_2 M-CAT gene conjugate with β -mercaptoethanol, and chromatographing the resulting mixture over Sepharose 4B. We discovered that the protein-specific 125I radioactivity previously associated with the $32P-$ labeled DNA was dramatically reduced by 75%. Taken together, these results strongly suggest that α_2M and the CAT gene are coupled by a disulfide bond.

Determination of the released 4-mercapto-2-nitrobenzoic acid at 412 nm showed that seven to nine lysyl residues on α ^M were modified. This modification should not affect its binding to cell surface receptors significantly. Previously a2M has been derivatized similarly and conjugated to horseradish peroxidase. The a2M-horseradish peroxidase conjugates were bound to cellular receptors and internalized by fibroblastic cells as well as unmodified α_2M (24). Furthermore, preliminary experiments showed that the internalization of the ∞ M-DNA conjugate by Swiss 3T3-4 cells is at least partially via the α ₂M receptor (unpublished results).

DISCUSSION

We describe here a simple method performed under mild conditions by which DNA can be conjugated to proteins. The method has been designed so that the conjugation takes place at unpaired quanine residues that have been introduced at the ends of the DNA. This should not interfere with the coding capacity of the DNA. Such conjugates should have many uses. One that we are interested in is to use such conjugates to direct DNA to recipient cells having specific receptors or other recognition components on their surface. Because many tissue culture cell lines have a₂M receptors on their surface, an a₂M-DNA conjugate containing either the thymidine kinase gene or the CAT gene should be useful in initiating such investigations. Either of these genes could be effectively used as markers for expression. Another possible use concerns the nature of DNA replication. A number of viruses have been found to have proteins conjugated to DNA in their virions. Such complexes have been suggested to have a role in DNA replication (27). The ability to synthesize DNA-protein conjugates may facilitate understanding of the role of such complexes.

ACKNOWLEDGMENTS

We wish to thank C. Gorman and B. Howard for the plasmid pASV-cat. G.T. Merlino was supported by a postdoctoral fellowship from the Arthritis Foundation.

REFERENCES

- 1. Sperling, J. and Sperling, R. (1978) Nucl. Acids Res. 5, 2755-2773.
2. Fiser, J., Margaritella, P. and Küchler, E. (1975) FEBS Lett. 52. 2
- 2. Fiser, J., Margaritella, P. and KUchler, E. (1975) FEBS Lett. 52, 281-283.
- 3. Rinke, J., Yuki, A. and Brimacombe, R. (1976) Eur. J. Biochem. 64, 77-89.
- 4. Baumert, H.G., Skold, S.E. and Kurland, C.G. (1978) Eur. J. Biochem. 89, 353-359.
- 5. Oste, C. and Brimacombe, R. (1979) Mol. Gen. Genet. 168, 81-86.
- 6. Vanin, E.F., Burkhard, S.J. and Kaiser, I.I. (1981) FEBS Lett., 124, 89-92.
- 7. Rinke, J., Meinke, M. and Brimacombe, R. (1980) J. Mol. Biol. 137, 301-314. 8. Millon, R., Olomucki, M., Le Gall, J.Y., Golinska, B., Ebel, J.P. and Ehresmann, B. (1980) Eur. J. Biochem. 110, 485-492.
-
- 9. Staehelin, M. (1959) Biochim. Biophys. Acta 31, 448-454. 10. Shapiro, R. and Hachmann, J. (1966) Biochemistry 6, 2799-2807.
- 11. Noller, H.F. and Chaires, J.B. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 3115-3118.
- 12. Noller, H.F. (1974) Biochemistry 13, 4694-4703.
- 13. Noller, H.F. and Herr, W. (1974) J. Mol. Biol. 90, 181-184.
- 14. Chapman, N.M. and Noller, H.F. (1977) J. Mol. Biol. 109, 131-149.
- 15. Hogan, J.J. and Noller, H.F. (1978) Biochemistry 17, 587-593.
- 16. Expert-Beganqon, A. and Hayer, D. (1980) Eur. J. Biochem. 103, 365-375.
- 17. Enquist, L.W., Vande Woude, G.F., Wagner, M., Smiley, J.R., and Summers, W.C. (1979) Gene 7, 335-342.
- 18. Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982) Molec. Cell. Biol. 2, 1044-1051.
- 19. Rabjohn, N. (1949) Organic Reactions, vol. 5, pp. 331-386, John Wiley, New York.
- 20. Tiffany, B.D., Wright, J.B., Moffett, R.V., Heinzelman, R.V., Strub, R.E., Aspergren, B.D., Lincoln, E.W. and White, J.L. (1957) J. Am. Chem. Soc. 79, 1682-1690.
- 21. Willingham, M.C., Maxfield, F.R. and Pastan, I.H. (1979) J. Cell. Biol. 82, 614-625.
- 22. Wickerhauser, M. and Hao, Y.L. (1972) Vox Sanguinis 23, 119-125.
- Mosher, D.F., Saksela, O. and Vaheri, A. (1977) J. Clin. Invest. 60, 1036-1045.
- 24. Dickson, R.B., Nicolas, J.C., Willingham, M.C. and Pastan, I.H. (1981) Exp. Cell. Res. 132, 488-493.
- 25. Habeeb, A.F.S.A. (1972) in Methods in Enzymology, Hirs, C.H.W. and Timasheff, S.N. Eds., Vol. 25, pp. 457-468, Academic Press, New York.
- 26. Roychoudhury, R., Jay, E. and Wu, R. (1976) Nucl. Acids Res. 3, 863-877.
- 27. Salas, M. and Vinuela, E. (1980) Trends in Biochemical Sciences 5, 191-193.