GENE TRANSFER BY RECEPTOR-MEDIATED ENDOCYTOSIS: STABLE EXPRESSION OF *NEO* FOLLOWING INSULIN-DIRECTED ENTRY INTO HepG2 CELLS

Barbara Isobel Huckett

Submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy
in the
Department of Biology
University of Natal
1989

Durban
1989

ABSTRACT

Evidence is presented for DNA delivery to cultured HepG2 hepatoma cells via the endocytotic pathway, under the direction of insulin, in a soluble system of transfection leading to stable gene expression. Serum albumin treated at pH 5.5 and 20°C for 48-60h with the water-soluble carbodiimide *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride has been found to produce positively charged *N*-acylurea albumin capable of binding different types of DNA in a reaction which is at least partially electrostatic in nature (Huckett *et al*, 1986). *N*-Acylurea albumin, synthesised at an albumin to carbodiimide mole ratio of 1:500, resulting in the attachment of 27 *N*-acylurea moieties per albumin molecule, was covalently conjugated to insulin by glutaraldehyde cross-linkage in order to produce a macromolecule, insulin-[*N*-acylurea albumin], with the facilities for both DNA transport and receptor binding. The resultant conjugate, purified by gel filtration through Sephadex G-100, was characterised in terms of molecular size, charge properties and insulin content by polyacrylamide gel electrophoresis, agarose gel electrophoresis and immuno-dot-blotting respectively. The conjugated protein was shown by gel band shift and nitrocellulose filter binding assays to bind DNA non-specifically in a reversible reaction which occurs rapidly, is dependent upon protein concentration and the ionic strength of the medium, and involves at least two types of intermolecular interaction. Furthermore, the conjugate was shown by

competitive displacement of [^{125}I]insulin to bind specifically and particularly avidly to the HepG2 insulin receptor. When the expression vectors ptkNEO and pAL-8 which incorporate the *neo* gene were complexed to the conjugate in an *in vitro* transfection procedure using HepG2 cells, G418 resistant clones developed at frequencies of 2.0 - 5.5 X 10^{-5}, possibly dependent upon vector promoter. Subsequently, a 923bp PstI fragment within the *neo* sequence was identified by Southern transfer in genomic DNA extracted from transfected cell populations which had been grown on a G418 regime through several subculture passages over a period of 44 days.

PREFACE

The experimental work described in this thesis was conducted in the Department of Biochemistry, University of Durban-Westville, Durban, under the joint supervision of Professor Arthur O. Hawtrey and Professor Mario Ariatti.

The studies embodied herein represent original work carried out by the author, and have not been submitted in alternative form to any other University. The work of other individuals incorporated into the thesis is fully acknowledged in the text.

ACKNOWLEDGEMENTS

My primary appreciation and thanks go to Professor Arthur Hawtrey and Professor Mario Ariatti, who provided the initial suggestion and subsequent persuasion that a part of the research work I was involved in at the University of Durban-Westville should be utilised for higher degree purposes. I am indebted to both of them for their consistently enthusiastic guidance and support in the work itself, as well as its channelling towards this thesis. I am grateful for the quality of their analytical and intuitive grasp of chemistry and biochemistry, their ability to engender interest and discussion in a wide variety of chemical and biological topics, and their willingness to teach laboratory skills, old and new, with meticulousness and the benefit of many years' experience. I thank them in particular for their personal interest and readiness to help at all times.

A number of other individuals contributed to the work presented. Professor Peter Jones of the Comprehensive Cancer Center, University of Southern California School of Medicine, donated the expression vector pAL-8. Dr Shirley Taylor of the Division of Biology, California Institute of Technology, provided the expression vector ptkNEO, which was constructed by Dr Barbara Wold of the same institution. Professor Dave R. Woods of the Microbial Genetics Research Unit, University of Cape Town, supplied *Escherichia coli* strain HB101. Dr Gerry Coetzee of the Department of Medical Biochemistry, University of Cape Town, provided the HepG2 cell line, and Dr Denver Hendricks of the Department of Biochemistry, University of the Western Cape, gave valuable recommendations concerning its routine maintenance. Mrs Lorraine van Hooff, Head of the Faculty of Science Photographic Unit, University of Durban-Westville, gave expert advice and assistance with a number of aspects of photography. Dr June Lawton, Head of the University of Durban-Westville Electron Microscope Unit, provided the use of an osmometer. Mr Bipraj Singh, Department of Virology, University of Natal Medical School, conducted all the gamma counting. Facilities for the laser printing of the document were made available by the University of Durban-Westville, and both Mr Graham Lundie of the Computer Centre and Mr Carl Juta of Computer Services gave valuable assistance in this regard. I thank each of these people for their involvement, interest and generosity.

Members of staff in the Department of Biochemistry, University of Durban-Westville, have contributed indirectly to the thesis by providing companionship and useful discussion at the laboratory bench at various times since the work began in 1984. In particular I would like to mention Dr Zainub Khan, Hasha Gordhan, Nirinjini Naidoo, Richard Hawtrey and Solly Weiler. In addition, there are colleagues in the Department of Biology, University of Natal, Durban, as well as ex-students of the same Department and a number of personal friends who have offered support in innumerable ways during the past seventeen years. I thank them all sincerely.

I acknowledge with special thanks the assistance afforded by Mr Colin Sapsford of the Department of Biology, University of Natal, in acting as administrative supervisor of this thesis.

The work presented was jointly supported by the Foundation for Research Development and University of Durban-Westville Research Funds.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 DNA UPTAKE BY LIVING CELLS - AN HISTORICAL PERSPECTIVE

1.1.1 *GENETIC TRANSFORMATION AS A NATURAL PHENOMENON IN BACTERIA*

In 1928, Griffith published the now classic observation that cells of *Diplococcus pneumoniae* could be heritably altered by association with a cell-derived "transforming principle". Sixteen years later this transforming substance having the potential for genetic interaction was identified as deoxyribonucleic acid, DNA (Avery *et al*, 1944). Within the next three decades the phenomenon of bacterial transformation - recognised as the uptake and integration into a recipient cell chromosome of DNA fragments released into the environment by death of a donor bacterium - had become widely identified as a natural mechanism of genetic recombination in certain eubacterial genera (Hotchkiss and Gabor, 1970; Notani and Setlow, 1974). During the same period the phenomenon came to be exploited extensively as a means of genetic analysis, especially in the important Gram positive model bacterium *Bacillus subtilis* lacking conjugation and transduction mechanisms (Dubnau *et al*, 1967). Techniques were developed to enhance transformation in responsive bacterial types. Commonly, nutrients in the growth medium were minimised, and the cell population allowed to reach a state of semi-starvation. This kind of treatment was found to induce in the cells a particular state of receptivity to DNA termed *competence*, which was later recognised as a protein-signalled event involving the production of specific cell surface receptors for the binding of DNA and its transport to

the cell interior (Mandelstam *et al*, 1982).

1.1.2 *CIRCUMVENTION OF NATURAL BARRIERS TO TRANSFORMATIONAL CHANGE IN BACTERIA*

By the early 1970s transformation had been shown to occur naturally in prokaryotes other than eubacteria. For example, it was demonstrated in the cyanobacterium (blue-green alga) *Anacystis nidulans* by Shestakov and Khyen (1970) and in the actinomycete *Thermoactinomyces vulgaris* by Hopwood and Wright (1972). Furthermore, more extreme methods were being developed at that time to provide conditions conducive to transformation in organisms found earlier to be completely resistant to the process. Ironically, included among those previously untransformable types was the Gram negative model bacterium *Escherichia coli*. However, in 1970, Mandel and Higa reported that *E. coli* cells pretreated with calcium chloride at low temperature became more permeable and were capable of taking up both linear and circular forms of bacteriophage DNA.* Later, using a similar method, Cohen *et al* (1972) demonstrated the uptake of circularised bacterial plasmid DNA by *E. coli*.

Events such as these caused a shift in the definition of transformation to include the cellular entry and expression of DNA regardless of whether chromosomal integration and recombination followed. They also set a precedent for a spate of technical advances in transformation technology. Numerous variations on the calcium chloride treatment approach were published, many of them directed towards improving the efficiency of transformation for a particular bacterial strain. In addition, high transformation efficiency mutant strains (usually nuclease deficient) were sought and isolated -

---

* Note: The uptake of naked viral DNA by a cell was regarded as a specific category of transformation and was termed *transfection*.

especially in *E. coli* - for application with calcium chloride procedures. Many such methods are still in common usage, particularly for the induction of recombinant plasmid uptake (Maniatis *et al*, 1982; Hanahan,1985) (section 1.1.4).

1.1.3 *GENETIC TRANSFORMATION IN EUKARYOTES*

As interest in bacterial transformation developed, increasing numbers of workers turned their attention to the possibility of naked DNA mediating comparable transformations in higher organisms.

Much of the early work was both ambitious and crude. It involved the injection of DNA from the tissue of one animal into the body of a whole animal of the same species but different phenotype. The pioneering work in this area of experimentation is that of Benoit *et al* (1957) who noted significant physical alterations after treating Pekin ducks with DNA extracted from those of the Kaki Campbell type. Working along the same lines, Martinovitch *et al* (1961) obtained changes in feather colour associated with gross morphological abnormalities following the injection of chicken embryos with DNA administered via the venous system. The chicken embryo experiments were followed up with observations of individuals of the next generation, and it was noted that in addition to the colour modifications they carried an apparently novel type of haemoglobin (Martinovitch *et al*, 1962). A degree of reproducibility was achieved in this particular work, and numerous modified offspring were obtained as a result (Martinovitch *et al*, 1961). Possible confirmation of the transformations shown by Martinovitch was provided by various reports of the appearance of feather colour changes as a consequence of whole blood injections from one type of fowl into another (Kushner, 1957; Stroun *et al*, 1962). Since avian erythrocytes are nucleated those experimental results are likely to be indicative of DNA interaction. Further in this regard, Leroy *et al* (1964) noted alterations in feather pigmentation

when guinea fowl erythrocyte nuclei were lysed and injected into Rhode Island Red fowls.

In general, however, attempts to reproduce transformations using gross DNA treatments such as these met with a high level of failure (Benoit, 1960) and a large number of papers were published to this effect. Attempts to extend the transformation phenomenon to other animal taxa were also fairly unsuccessful (Billet *et al*, 1964).

Considerably more elegant experiments were set up by workers who turned to cell culture systems to look for transformation in terms of the transfer and expression of specific genetic markers. Two examples, both involving mammalian cell culture, will illustrate this point. Kraus (1961) demonstrated that bone marrow cells taken from sickle cell anaemia patients which lack the ability to produce the β^A chain of haemoglobin may have that ability conferred upon them by incubation with DNA from normal bone marrow cells homozygous for haemoglobin. In experiments of parallel design, Szybalska and Szybalski (1962) grew mutants of a human cell line deficient in the enzyme inosinate pyrophosphorylase and treated them with extracts of DNA prepared from enzyme-positive cells of the same line. The appearance of enzyme-positive genetically transformed cells resulted from this procedure, transformation occurring at a frequency of 1 event per 4×10^4 cells in the presence of 150µg DNA/ml.

The pathway and mechanism of DNA uptake in successful transformations such as those quoted above was unknown, although attempts had already been made and continued to be made to follow the fate of DNA from the exterior environment, across the cell membrane, into the interior of the cell, and even from organ to organ. For example, Gartler (1959) showed that Earle L cells took up a small amount of [^{3}H]-homologous DNA in intact polymerised form, with a smaller amount being degraded by the cells. The same results were obtained when DNA labelled with bromodeoxyuridine was used (Gartler, 1960). Sirotnak and

Hutchinson (1959) used [^{32}P]-labelled DNA to show that its rate of uptake by L1210 cells growing intraperitoneally was very rapid, maximal in 3 minutes, and Tsumita and Iwanaga (1963) demonstrated that 80-90% of [^{3}H]-DNA injected into mice intravenously disappeared from the blood stream in 10 minutes compared with only 20% of hydrolysed [^{3}H]-DNA in the same period. The fate of [^{3}H]-DNA in germinating barley seeds was studied extensively by Ledoux and coworkers. In one series of experiments it was shown that a 30 minute incubation of the endosperm region of the seedlings with labelled DNA was sufficient for the DNA to become localised in the nuclei of the root cells, some distance away (Ledoux, 1965).

In reviewing the field of DNA uptake by living cells, Ledoux (1965) offers speculations on mechanism which appear especially insightful from present day perspective. At a time when modern cell biology was in its infancy, and before receptor-mediated endocytosis was a recognised phenomenon, he suggested the importance of a kind of indiscriminate pinocytotic process in which selectivity of molecular uptake could be provided by specific site attachment in the cytoplasm. At the same time he did question whether transformation is, in fact, a natural process in eukaryotes.

In concluding his review, Ledoux states:

> "We have already good indication of the possibility of DNA uptake by living cells; there seems to be little doubt concerning the existence of such a process, at least in certain cell types. Many positive results have already been obtained for cells growing in suspension....
> ....concerning DNA uptake by tissues and organs, we are on much less certain ground, not because contradictory results have been obtained, but mainly because the results are still fragmentary and the results not numerous."

His final comment is particularly perspicacious:

> "It also appears more and more evident that the isolated cases of transformation reported in the literature should not be considered too lightly as they could indicate the appearance of a new dimension in molecular biology."

By the end of the 1960s the evidence for DNA uptake by eukaryotic cells had been significantly consolidated by the publication of a number of detailed studies (Ayad and Fox, 1968; Ledoux and Huart, 1968,1969; Laval *et al*, 1970). Then, in 1971, a piece of work by Merril and coworkers emerged in which it was reported that the specific bacterial gene sequence coding for α-D-galactose-1-phosphate uridyl transferase, incorporated into the DNA of transducing phage Lambda, could be taken up and expressed by human fibroblasts isolated from a patient with congenital lack of the enzyme. This work was seen to be of key significance, not only because the incorporated gene was bacterial in origin yet stably inherited and expressed in a mammalian system, but also because the gene was provided to the system in purified form in multiple copies for the first time. The paper was the subject of a major editorial comment in the issue of *Nature* in which it was published. The following phrases from the editorial reflect the importance put on the work:

> "The experiments reported by Merril, Geier and Petricciana....will be of the greatest and most startling interest to all biologists....they will have very far reaching implications indeed."

Those "far reaching implications" began to be realised almost immediately, for as the type of work reported by Merril's group developed, it met and ultimately amalgamated with radical new technical advances in nucleic acid chemistry and molecular biology which, in turn, allowed exponential growth of the field in a quite unprecedented manner.

1.1.4 *DEVELOPMENT OF THE NEW RECOMBINANT DNA TECHNOLOGY*

Advances in molecular biology from the mid-1970s to the present time represent the growth of a technology of dramatic proportions and have led to a considerable information explosion in the field of biochemical genetics. This has come about largely because a number of individually powerful techniques has

been integrated to form an array of compatible tools having the potential to tackle and solve a wide variety of gene transfer problems. Key components in this manifold instrumentarium include

- enzymes for the precise fragmentation and ligation of DNA molecules;
- methods for the separation and sequence identification of DNA fragments, and
- a variety of nucleic acid vectors, viral or plasmid in origin, into which DNA fragments of interest can be inserted accurately and which can be replicated in a host cell, thereby *cloning* the introduced sequences of foreign DNA.

Manipulation of the DNA during the overall cloning process takes place alternately within and without cell structures. The original viral or plasmid vector structure is produced in cultured cells, then extracted and biochemically modified by the insertion of appropriate DNA sequences before reintroduction into a host cell for replicative proliferation. The host cell used at this stage is rarely the exact cell type of origin; in the case of bacterial plasmids, the most versatile of vectors, the cloning is usually implemented in a well characterised bacterial host, the one most often chosen being "that workhorse of Molecular Biologists, *E. coli*" (Glover, 1985). Introduction of the newly designed plasmid into the host is by transformation. Thus the transformation methodology discussed above in section 1.1.2 is yet another device in the overall repertoire of techniques, and one of considerable importance. Subsequent to this cellular step the cloned plasmid, now in considerable quantity, is retrieved by extraction for experimental use.

It is this controlled recombination of known and defined DNA sequences *in vitro* together with the production of multiple copies of the recombinant DNA *in vivo* which gives the new technology its powerful experimental thrust.

DNA sequences which lend themselves to cloning include control regions as well as structural genes *per se*. Moreover, these may be either prokaryotic or eukaryotic in nature, even though carried on plasmid structures and replicated in bacteria. It is therefore not surprising that, as recombinant DNA technology has evolved, the prospect of specific gene transfer and expression in eukaryotic cell systems has become a realistic one. Currently, major interest is being shown in this area of manipulative genetic practice, and numerous approaches are being developed with varying degrees of applicability and success. Some of the central aspects of present day gene transfer technology as applied to eukaryotic systems are reviewed below in section 1.2.

1.2 CURRENT STATUS OF EXPERIMENTAL GENE TRANSFER TO EUKARYOTES

1.2.1 *PLASMID VECTORS USED IN EUKARYOTIC CELL EXPRESSION*

Plasmid vectors designed for gene transfer and expression in eukaryotic cells must fulfil a number of functions and are constructed accordingly. Firstly, the plasmid must carry an appropriate replication origin in order to be capable of propagation in a bacterial host. Secondly, there must be present on the plasmid a selective marker such as a gene coding for antibiotic resistance as a means of plasmid maintenance during bacterial host replication. Both of the aforementioned characteristics are features of some naturally occurring plasmids. A third component, a suitable "test" gene, often called a reporter gene, must be built into the plasmid structure for the monitoring of gene expression in the recipient eukaryotic cells. While the test gene itself may be either prokaryotic or eukaryotic in origin, the fourth component, consisting of promoter and possibly also enhancer elements for the initiation and control of test gene transcription, is necessarily eukaryotic in action and is taken from a viral

genome in many instances. This promoter sequence is usually inserted upstream from the test gene as an adjacent flanking region. A fifth functional sequence involved in the processing of the transcripts must be added to the plasmid structure in order to ensure that efficient translation takes place. Such a sequence includes, for example, a section controlling polyadenylation of mRNA, and is placed so that it flanks the test gene in a downstream position. It is frequently taken from the same viral source as the promoter region. Introns for the control of mRNA splicing may be included in this portion of the plasmid as well.

Many eukaryotic expression vectors are based on the plasmid construct pBR322 (Bolivar *et al*, 1977) which bears an origin of replication from pMB1, similar to that of the natural *E. coli* plasmid colE1-K30, as well as the gene for ampicillin resistance (*amp*^{R}) coding for a ß-lactamase which is part of the transposon Tn*3* from the *Salmonella paratyphi* plasmid R1, and genes coding for tetracycline resistance (*tet*^{R}) taken from the *Salmonella panama* plasmid pSC101. Inactivation of one of the two antibiotic resistance gene sequences by insertion of foreign DNA provides an excellent means for insertion selection. In order to increase the usefulness of pBR322 further in this respect, Bolivar and coworkers have deleted a number of endonuclease-sensitive sites so that a large number of restriction enzymes cut the plasmid only once and a small number cut it only twice. An additional advantage of using pBR322 is that its nucleotide sequence has been fully determined (Sutcliffe, 1979). One much used family of expression vectors incorporating the origin of replication and *amp*^{R} from pBR322 is the prototype pSV_2 and its many derivatives, in which the control sequences are provided by Simian Virus 40 (SV40) (Mulligan and Berg, 1980).

There are two fundamental types of expression vector in use in eukaryotic cells. The first type is designed and used to estimate transient expression shortly after uptake, while the second type is suited for the detection of expression following

stable integration into the recipent cell genome. Short term expression as demonstrated by the former is often used to estimate important experimental parameters such as the efficacy of a particular promoter in the cell line under investigation as a preliminary to applying the latter.

Transient expression vectors incorporate test genes coding for products distinguishable experimentally - as either RNA or protein - from those of the recipient cell. Activity of the test gene can be assayed, therefore, within hours of entry, at a stage when the plasmid vectors are not yet integrated into the host cell chromatin. At this point the plasmids which have been assimilated are often linked into large concatemers termed transgenomes (Scangos and Ruddle, 1981) and the level of expression is very high. However, transgenomes lack functional centromeres and are soon lost from the cell population, so that the expression peaks 24 - 72 hours after uptake of the DNA. Specialised genes from bacterial sources are well suited to use as reporters, since the recipient eukaryotic organisms lack the gene and provide a null background against which the assay can be conducted. Two which are commonly used are *cat* and *lacZ* which respectively encode the enzymes chloramphenicol acetyltransferase (Gorman *et al*, 1982) and ß-galactosidase (Hall *et al*, 1983). A test gene which has become much used in plant systems recently (Jefferson *et al*, 1987) is the ß-glucuronidase (GUS) coding sequence from the *uidA* locus of *E. coli*; the enzyme causes hydrolytic cleavage of a wide variety of ß-glucuronides of which many are available commercially as spectrophotometric, fluorometric and histochemical substrates.

By contrast, *stable expression vectors* carry genetic markers whose products, albeit generated at low levels, allow selective survival of the recipient cells in culture. The test of their presence is thus a relatively long term process and integration of plasmid into the genetic apparatus of the recipient cell is necessary so that continuous *de novo* synthesis of gene products occurs. The eukaryotic *tk* gene coding for the enzyme thymidine

kinase (TK) has been used a great deal as a test gene in stable expression vectors, one of the first such vectors to be designed being based on the pBR322 construct (Enquist *et al*, 1979). The selection system, invented by Szybalska and Szybalski (1962) and developed by Littlefield (1964) for application in early genetic studies (section 1.1.3), requires the use of HAT culture medium, which contains hypoxanthine, aminopterin and thymidine. In this medium aminopterin blocks the pathway leading to the essential purine intermediate inosine monophosphate (IMP) as well as the conversion by methylation of the pyrimidine dUMP to dTMP, thereby creating dependence upon extraneous sources of purines and thymidine. The hypoxanthine included in the medium can be used as a source of purines since it is phosphorylated by the enzyme hypoxanthine-guanine phosphoribosyl transferase to form IMP. Similarly, the thymidine present can be phosphorylated by the TK enzyme. As a result, cells which are tk^+ survive in the HAT medium while those which are tk^- do not. A disadvantage of the system is that the selective advantage of *tk* uptake can only be seen in recipient cells which are mutants deficient in the expression of TK (Wigler *et al*, 1977).

By comparison with *tk*, the bacterial gene neo^R is a far more universally applicable test of stable integration as it is a dominant marker which does not require a genetically deficient recipient cell for expression to be observable (Colbere-Garapin *et al*, 1981). Selection using neo^R can be applied, therefore, to almost any cell type. The gene neo^R, borne in the natural state on the transposon Tn*5*, codes for the aminoglycoside 3′-phosphotransferase Type II (APH(3′)-II) and confers resistance to the antibiotic neomycin upon bacterial cells which carry it. Neomycin itself interferes with the action of prokaryotic but not eukaryotic ribosomes, but an analogue to it, G418 (Geneticin), is effective against eukaryotic animal ribosomes while still being inactivated by the APH(3′)-II enzyme. Clearly, then, animal cells bearing and expressing the neo^R gene can survive in culture medium containing inimical levels of G418. Comparable selection of plant tissues carrying the neo^R

gene can be made using a growth regime which incorporates kanamycin (Rogers *et al*, 1986).

1.2.2 *CURRENT METHODS FOR THE INTRODUCTION OF DNA INTO EUKARYOTIC CELLS*

The methods utilised for the transfer of genetic material into eukaryotic cells at the present time are extremely varied. Almost without exception, however, they involve either the active insertion of DNA into cells or the active uptake of DNA by cells via specific mechanisms that are at least partially understood. This is in contrast with the apparently passive cellular incorporation of DNA by unknown means under conditions of straightforward incubation as described by workers such as Szybalska and Szybalski (1962) and Merril *et al* (1971) referred to in section 1.1.3. Some of the current methods are simple in concept; others are based on highly contrived strategies. Overall, the appeal of these techniques lies in the relatively high frequencies of gene transfer, genomic integration and ultimate expression which result from their application, and the consequently good reproducibility levels.

During the course of development of this field of work, common use of viral DNAs in early transfer studies led to adoption of the term *transfection* to describe the process of gene delivery to eukaryotic cells, and today this term may be used whether the DNA being transferred is viral in origin or not. While the term *transformation* is still used validly for the same process and the phenotypic expression which results from it, it is likely that those workers who avoid its use in favour of *transfection* do so in order to avoid confusion with the alternative and highly specific meanings attached to the word *transformation* in the fields of bacterial genetics and oncology respectively. This thesis draws upon a wide range of work in which, variously, each of the two terms is used to describe the introduction of DNA into eukaryotic cells; therefore *transfection* and

transformation are used interchangeably to convey this meaning in the sections which follow.

Two broad categories of transfection methodology exist. In the first, the DNA of interest is introduced to the cells in the unprotected state (naked DNA transfer). In the second, the DNA is packaged in some way inside a membranous coating of non-nucleic acid material prior to cell contact (vehicle-mediated DNA transfer).

Naked DNA Transfer

An extensive range of methods exists for transfer to recipient cells of DNA which is unprotected. Despite the potential for exposure of the DNA to a variety of biochemical hazards during the transfection process, these methods are the most commonly used, probably because they are relatively simple. In many instances the DNA is in solution, and it is the solution which is delivered into the cell interior. In other cases the DNA is precipitated to form a granular suspension to which recipient cells are exposed, or the precipitate is taken a step further and coated on to particulate support material prior to solid phase delivery.

Micropuncture

Microinjection. The first injection of substances into cells using glass micropipettes was reported early in this century (Barber, 1914), and in a classic piece of work some years later the technique was used under darkfield illumination with cells in tissue culture (Chambers and Fell, 1931). In the modern version of the method, a glass microcapillary with a tip 0.2 - 1.0μm in diameter is directed spatially by a mechanical micromanipulator during phase contrast microscopic observation of the cells to be injected, and samples of 1 - 2 X 10^{-11}ml are transferred to the cells by positive and negative air pressure

control exerted by a syringe connected to the capillary (Diacumakos, 1980; Graessman and Graessman, 1983). Up to 1000 cells can be injected per hour (Mueller *et al*, 1980). The degree of precision is such that injected material can be placed in specific subcellular locations within the recipient cell, including the nucleus (Diacumacos *et al*, 1970; Diacumacos, 1980). Puncturing of the nuclear membrane does not lead to cell death (Diacumakos, 1980) and, in general, cultured cells are affected only mildly, if at all, by the intrusion of the syringe and the introduction of modest volumes of material into the intracellular space (Stacey, 1980).

Many microinjection studies have dwelt upon the effects of inserting specific proteins and RNA species into cells. However, microinjection of primary genetic material into cells has been exploited extensively and has included the transplantation of nuclei (Graessman, 1970) and chromosomes (Diacumakos *et al*, 1970), as well as the transfer of purified DNA (Graessman and Graessman, 1976; Palmiter *et al*, 1982). Microinjected DNA species are efficiently expressed; following intranuclear injection, transient expression levels are 50 - 100%, and significant stable expression as shown by 0.1 - 30% transformed colonies is achieved (Graessman and Graessman, 1983; Celis, 1984).

As Diacumakos (1980) has pointed out, the advantages of the method include its directness - the immediate examination and monitoring of treated cells - and the requirements for few cells and small amounts of material to be transferred. Graessman and Graessman (1983) consider that a distinctly attractive aspect of the method is that opportunity is open to the investigator to choose and vary the site of inoculation within the cell. They make the additional comment that every tissue culture cell line tested has proven suitable for microinjection, including suspension cultures, as these can be bound to a substrate for the duration of the procedure. In spite of these positive features, the method has been criticised on the grounds that it

is laborious (Loyter *et al*, 1975), requires particular skill (Tao *et al*, 1987) and is applicable to only small numbers of cells, thereby limiting the types of analysis which can be employed within the system (Celis *et al*, 1980). These various features of the microinjection approach have been reviewed and evaluated by Celis (1984) in relation to a number of other transfer methodologies. Clearly the strongest applications of the microinjection technique lie in precise studies of single cells or small populations of cells which lend themselves to the use of highly sensitive assays such as those employing immunofluorescent, autoradiographic, histochemical or virus plaque measurements.

Laser Micropuncture. Tao *et al* (1987) have described a technique of direct gene transfer into cultured cells using laser microsurgical procedures in an ultramodern variation of the microinjection method. As in manual microinjection, the cells can be visualised and the perforation directed, although the entry of DNA and its positioning within the cell cannot be precisely controlled. The suggested mechanism of action is that of puncturing of the cell surface followed by rapid self-sealing of the hole, with DNA being taken up from the medium in the intervening period, presumably by a process of diffusion.

The third harmonic 355nm wavelength from a short-pulsed Quantel model YG481a neodymium-ytrium-aluminium garnet laser was focused on the surface of each cell to be treated, and single laser pulses of 10nsec duration obtained by electronic shutter control. Pulse energies were in the range 23-67μJ. Cells were observed and the laser beam targeted by employing a motorised microscope stage coordinated with a video camera, TV monitor and computer. One thousand cells were irradiated per hour - a treatment rate comparable to that of experienced workers using the manual microinjection method (Mueller *et al*, 1980). Using human HT1080-6TG fibrosarcoma cells and the plasmid pSV_2neo as a model system, between 1 in 8×10^4 and 1 in 3×10^3 cells demonstrated expression of the *neo* gene at a plasmid DNA

concentration of 12μg/ml.

The authors make the point that although the manual microinjection method yields higher frequencies of DNA uptake and stable transformation in mammalian cells, it has worked only poorly in isolated plant protoplasts. They suggest that the laser method could be of importance in the genetic manipulation of plants and other types of cells for which similar difficulties have been encountered previously.

Electroporation. Electroporation - sometimes termed electropermeabilization - is the reversible electrical breakdown ("poration") of cell membranes brought about by high voltage electrical pulses (Neumann and Rosenheck, 1972; Serpersu *et al*, 1985). When cells are electroporated in the presence of DNA, its transfer into the cells is actively facilitated via a mechanism which is now known to be at least partially electrophoretic (Winterbourne *et al*, 1988). The production of holes, while quantitatively controllable, is random and therefore lacks the spatial precision of the microinjection procedure; on the other hand, large numbers of cells can be punctured and DNA-treated simultaneously. The method has been applied to a wide variety of cell types including mammalian cell lines (Potter *et al*, 1984; Stopper *et al*, 1987), yeast sphaeroplasts (Karube *et al*, 1985), plant protoplasts (Fromm *et al*, 1986) and intact prokaryotes (Taketo, 1988). There have been recent reviews on specific applications (Shigekawa and Dower, 1988; Morikawa *et al*, 1988).

Cells are treated in a cooled discharge chamber, 1 - 3ml in volume, fitted with twin parallel platinum electrodes connected to a pulse generator. During pulse treatment the cells are suspended in low ionic strength (low conductance) buffer made isotonic with nonionic molecules; after pulse treatment and prior to further culture, cells are transferred to a higher ionic strength medium containing magnesium and calcium ions for 20 minutes or so at 37°C to allow membrane resealing (Stopper *et*

al, 1987). In a typical procedure using mouse L cells and a plasmid carrying the selective marker gene *neo*, expression of the transferred gene was optimal when the cell concentration was 3 X 10^5/ml, the plasmid concentration 1µg/ml, the temperature 4°C, the field strength 10kV/cm, the duration of the pulse 5µsec, the number of pulses 10 and the interval between pulses 1 minute. These conditions yielded one *neo* transformant per thousand cells (Stopper *et al*, 1987).

A commercially produced apparatus for electroporation is now available which claims transfection efficiencies for mammalian cells even higher than those of Stopper and coworkers, although not for all cell types (Bio-Rad Laboratories, Bio-Radiations Bulletin No. 61EG, 1986). This apparatus is being used extensively as an alternative to calcium-induced plasmid transformation of bacteria; the resultant transfer efficiencies are greater (Taketo, 1988).

Perforation by High Velocity Microprojectiles. An innovative method for DNA delivery by micropuncture, namely the use of a particle gun capable of firing tungsten microprojectiles into plant cells through cellulosic cell walls, has been reported by Klein *et al* (1987). The gun is a simple barrel down which a cylindrical nylon projectile, 5mm in diameter and 8mm long, is accelerated by firing pin detonation of a gunpowder charge. The tungsten particles, 4µm in diameter, are suspended in 1 - 2µl of water and spread on the front surface of the nylon projectile. On detonation, the particles accelerate with the nylon projectile and continue moving towards the target cells through a 1mm aperture in a steel plate designed to stop the nylon projectile itself. The tungsten microprojectiles have an initial velocity of 430m/sec.

Cells in the form of 1cm^2 pieces of *Allium cepa* (onion) epidermis - approximately 2000 cells per piece - were used as targets and placed 10 - 15cm from the end of the firing device. When the tungsten particles in suspension were coated with

precipitates of either TMV RNA or p35S-CAT plasmid prior to bombardment, expression resulted.

In common with electroporation, the advantages of this method are its rapidity, resulting from the simultaneous delivery of genetic material to many cells, and the high frequencies of transformation observed. An additional point in its favour is that cell culture or cell suspension is unnecessary; whole tissues can be subjected to the procedure without pretreatment. The possibility exists that by using the technique on meristematic tissue or embryogenic callus, the problems attendant on genetically transforming graminaceous monocotyledonous plants might be circumvented, current approaches involving regeneration of whole plants from altered protoplasts or *Agrobacterium tumefasciens* infection being beset by serious limitations. The authors point out that for such development to be feasible the process must be refined so that smaller cells than those of onion epidermal tissue can be only just penetrated. More recent work demonstrates that a degree of sophistication has been achieved in this respect: the method has been applied in the transformation of chloroplast genomes in *Chlamydomonas* (Boynton *et al*, 1988) and mitochondrial genomes in *Saccharomyces* (Johnston *et al*, 1988) by the bombardment of monolayers of the respective cells types, both of which are smaller than those of onion epidermis.

It is interesting that Klein and coworkers make no comment on the use of animal tissue as target material. Suitably monolayered cells in whole tissue form are certainly not as commonplace in animals as they are in plants. It is also possible that the method, being particularly aggressive, would cause serious trauma in animal cells, although conditions of bombardment could perhaps be further modified to overcome such difficulties.

Membrane Disturbance by Mechanical Abrasion

One group of workers has developed the idea of loading macromolecules into mammalian cells in culture by mechanically inducing general wounding of the plasma membrane, thereby opening up the cell interior to normally impermeant substances. McNeil *et al* (1984) first reported this approach in a technique termed *scrape-loading* in which washed cells are scraped from the culture surface with a rubber policeman in the presence of the macromolecule of interest, resuspended further in culture medium, replated, then allowed to recover and spread. In an attempt to eliminate the cell suspension step with its concomitant requirement for replating and growth prior to further experimentation, the method has been refined in a variation called *bead-loading* (McNeil and Warder, 1987), which causes grazing of the upper surface of attached cells *in situ*.

In the bead-loading technique the culture medium of the cell monolayer is washed off and replaced by a small volume of a solution of the macromolecule to be loaded, glass beads 75 - 500μm in diameter sprinkled on to the cells, and the layer of beads caused to roll across the cells by rocking 3 - 6 times. After rinsing the beads and the exogenous molecules away, the bead-loading step is complete, and fresh medium is then added to allow growth to continue. The method is therefore exceedingly straightforward, requiring little expertise and only the most simple of tools. Moreover, variations in bead size and degree of agitation allow for adjustments to be made according to the fragility of the cells under study and the extent of loading desired. However, the authors have shown that increased abrasion and loading is always gained at the expense of cell viability.

Successful transformation of cells with genes borne on plasmids has been achieved using the scrape-loading method, with both transient and stable expression being noted (Fechheimer *et al*, 1987), but no reports are yet available indicating a similarly

successful use of bead-loading for DNA transfer.

Use of Chemical Facilitators

Polycations. When DNA is added to the culture medium in which mammalian cells are growing, its occasional uptake into the cytoplasm and nucleus (Kraus, 1961; Szybalska and Szybalski, 1962) is found to be enhanced to measurable levels in the presence of certain polymeric cationic compounds. Examples of such cationic facilitators are diethylaminoethyl-dextran (DEAE-dextran) (McCutchan and Pagano, 1968), polyornithine (Farber *et al*, 1975) and 1,5-dimethyl-1,5-diazaundecamethylene polymethobromide (Polybrene) (Isa and Shima, 1987). The mechanism of facilitation remains obscure, but seems likely to relate to complex ionic interactions in which the positively charged cations act as a bridge between the anionic nucleic acids and negative charges on the cell surface, thereby creating a physical link between the DNA and the membrane (Isa and Shima, 1987).

While cationic facilitators have been used widely to effect increased uptake of viral nucleic acid (McCutchan and Pagano, 1968), including supercoiled covalently closed circular forms (Sompayrac, 1981), transfer efficiencies have not been particularly impressive. Furthermore, in the case of selectable marker gene transfer, permanently altered cell lines have not been achieved using these methods (Gorman, 1985), although recent work by Isa and Shima (1987) suggests that DNA transfer followed by stable expression is possible using a cationic facilitator (Polybrene) in conjunction with dimethylsuphoxide (DMSO).

Calcium Phosphate. More successful chemically facilitated transformations by far have been demonstrated using calcium phosphate to precipitate DNA molecules on to mammalian cell surfaces. This technique was initially developed to detect the infective potential of viral DNA (Graham and van der Eb, 1973)

but has been used extensively since in a wide variety of mammalian genetic studies requiring the transfer of gene sequences from one cellular source to another.

The technique draws upon simple equipment, is applicable to most cells in monolayer culture and some cells in suspension, and can be used to produce either transient gene expression or stable transformation in a consistent manner. The fact that its efficacy does not depend upon membrane aggression, but rather leaves the cell surface intact, gives the method added appeal. On these grounds it has proved to be an extremely popular means of gene transfer to mammalian cells, and has been quoted as the most widely applied method (Vainstein *et al*, 1983; Gorman, 1985). The transfer frequencies obtained vary a great deal according to the recipient cell line. Tao (1987) quotes 1 - 5 transformants per 10^5 - 10^7 cells as a general value. One often quoted disadvantage lies in these only moderate frequencies (Strain and Wyllie, 1984; Tao, 1987; Tsao, 1987); another in the abnormally high levels of calcium and chemical adjuvants required. While this latter factor has in the past been the subject of some general suspicion and criticism (Tao, 1987), it has only recently been shown that calcium perturbation during the calcium phosphate procedure may lead to distinct genetic effects, namely heritable alterations in both the tumorigenic and malignant behaviour of transfected neoplastic cells (Kerbel *et al*, 1987) and abnormal transcriptional activation (Pine *et al*, 1988). Based on this evidence, use of the method may seriously decline, especially in studies where investigation of gene expression is the primary object.

The method involves the mixing of the DNA of interest with calcium chloride followed by gradual introduction of the mixture to phosphate ions in order to produce a fine coprecipitate of DNA and calcium phosphate. When this suspended precipitate is added to cell cultures it binds to the plasma membranes readily, and over the next several hours of incubation the DNA may be taken up into the cells and ultimately expressed. There are

many relevant factors to be considered in using the technique and consequently many versions of the basic protocol. Some of the more important parameters affecting efficiency have been examined and reviewed by Graham *et al* (1980). They include cell type, cell density, carrier DNA concentration and molecular weight, size of the coprecipitate particles and time of exposure of cells to particles. Two particular methodological variations subsequently proven to be the ones most commonly used have been described in detail by Gorman (1985). A further variation which utilises whole chromosomes rather than DNA fragments has been outlined by Srinivasan and Lewis (1980).

By employing fluorescent dye-stained DNA and tracing its pathway into the cell, Loyter *et al* (1982) have shown that the DNA-calcium phosphate complexes enter cells by an endocytosis- like process, and that the DNA, for the most part, becomes trapped within degradative vacuoles. Vainstein *et al* (1983) have argued that these findings explain respectively why cells such as mouse L cells with a high level of endocytic activity are well suited to transformation by DNA-calcium phosphate coprecipitates and why the observed transformation frequencies for all cells are never very high. In this regard it is significant that transformation levels are raised in the presence of chloroquine, a drug with lysosomotropic activity (Luthman and Magnusson, 1983).

Other additives which are observed to increase the frequency of gene uptake and expression after DNA-calcium phosphate treatment have been reported. For example, sodium butyrate has been shown to increase the expression of transferred DNA several-fold (Gorman *et al*, 1983), especially if the cells are shocked with 15% glycerol first (Gorman, 1985); the butyrate appears to enhance the transcriptional activity of plasmids bearing SV40, polyoma and papilloma virus control regions in particular. Addition of dimethylsulphoxide (DMSO) to cells at the end of the coprecipitate treatment period also has been shown to improve transfer and expression efficiency (Srinivasan and Lewis, 1980).

many relevant factors to be considered in using the technique and consequently many versions of the basic protocol. Some of the more important parameters affecting efficiency have been examined and reviewed by Graham *et al* (1980). They include cell type, cell density, carrier DNA concentration and molecular weight, size of the coprecipitate particles and time of exposure of cells to particles. Two particular methodological variations subsequently proven to be the ones most commonly used have been described in detail by Gorman (1985). A further variation which utilises whole chromosomes rather than DNA fragments has been outlined by Srinivasan and Lewis (1980).

By employing fluorescent dye-stained DNA and tracing its pathway into the cell, Loyter *et al* (1982) have shown that the DNA-calcium phosphate complexes enter cells by an endocytosis- like process, and that the DNA, for the most part, becomes trapped within degradative vacuoles. Vainstein *et al* (1983) have argued that these findings explain respectively why cells such as mouse L cells with a high level of endocytic activity are well suited to transformation by DNA-calcium phosphate coprecipitates and why the observed transformation frequencies for all cells are never very high. In this regard it is significant that transformation levels are raised in the presence of chloroquine, a drug with lysosomotropic activity (Luthman and Magnusson, 1983).

Other additives which are observed to increase the frequency of gene uptake and expression after DNA-calcium phosphate treatment have been reported. For example, sodium butyrate has been shown to increase the expression of transferred DNA several-fold (Gorman *et al*, 1983), especially if the cells are shocked with 15% glycerol first (Gorman, 1985); the butyrate appears to enhance the transcriptional activity of plasmids bearing SV40, polyoma and papilloma virus control regions in particular. Addition of dimethylsulphoxide (DMSO) to cells at the end of the coprecipitate treatment period also has been shown to improve transfer and expression efficiency (Srinivasan and Lewis, 1980).

(This effect of DMSO has been noted above in relation to polycation-induced DNA transfer.) In a detailed study of the uptake and stability of SV40 DNA after calcium phosphate transfection of CV-1 cells, Strain and Wyllie (1984) conclude that the mechanism of action of agents such as glycerol and DMSO remains enigmatic, but may involve the conferral of greater stability on the transferred DNA intracellularly rather than the mere alteration of its rate of entry into the cell.

Vehicle-Mediated DNA Transfer

A number of methods have been developed for the transfer to mammalian cells of DNA which is enclosed in a protective covering. In each of the methods the covering is a membrane of natural or artificial origin which offers distinct advantages to the DNA being carried. The membrane may serve the purpose of, firstly, enclosing large numbers of DNA molecules (gene copies) in each vehicular unit; secondly, protecting the DNA molecules from attack by serum or cell-secreted nucleases prior to cell entry, and thirdly, attaching to and either fusing with or being taken up by the target cell, via natural or contrived mechanisms, thereby allowing intracellular unloading of the DNA to take place in a manner which bypasses the normal permeability barrier of the cell membrane. The various experimental approaches employing membranous DNA carriers are presented below.

Protoplasts as Carriers

DNA present in bacterial cells, including multiple cloned sequences carried on plasmids, can be transferred to mammalian cells directly and very effectively if the cells are converted to wall-less protoplasts and the protoplasts brought into close contact with the recipient cells by fusion techniques. The charm of the method lies in the fact that the DNA of interest never leaves its natural cytoplasmic environment and therefore

does not have to be extracted or purified at any stage. This saves time and effort and, more importantly, eliminates procedures and conditions liable to cause nicking and shearing of the DNA.

Protoplast fusion was developed initially by Schaffner (1980) for the introduction into mammalian cells of SV40 DNA cloned in *E. coli*. The technique was extended by Sandri-Goldin *et al* (1981) using cloned DNA sequences from herpes simplex virus type 1 (HSV1), including the thymidine kinase (*tk*) gene, and these same authors subsequently reviewed the general methodology (Sandri-Goldin *et al*, 1983). More recent work reports significantly successful use of protoplast fusion as a method of choice in permanently transferring the genes *cat* and *neo*, incorporated into respective plasmids, to human haematopoietic cells (Tsao *et al*, 1987).

Three steps are involved in the procedure: the transformation of and subsequent production and amplification in an appropriate bacterial host of plasmids carrying the gene or genes of interest; the removal of cell wall material from the whole bacterial cells by lysozyme digestion in order to create protoplasts, and fusion of the resultant protoplasts to recipient eukaryotic cells in culture. The bacterial strains used as carriers in experiments involving mammalian cell recipients have been mostly mutants of *E. coli*- HB101, 1106 and DH-1 - while those employed as carriers in the genetic manipulation of higher plant cells have been variants of *Agrobacterium tumefasciens* (Sandri-Goldin *et al*, 1983). The recipient mammalian cell lines used have been many and varied, and have included 3T6 mouse fibroblast cells, Vero cells (from African Green Monkey kidney) and human HeLa cells (Schaffner, 1980; Sandri-Goldin *et al*, 1981) as well as human leukaemia K562 cells (Tsao *et al*, 1987). Plant work has included the use of cells from the drug-producing *Vinca rosea* as recipients (Hasezawa *et al*, 1981).

The actual binding process between the two cellular elements is brought about by pelleting the protoplasts on to recipient cell monolayers or by pelleting a mixture of protoplasts and cells, then overlayering the close-packed mixture with concentrated (approximately 50%w/v) polyethylene glycol (PEG) for 2 - 4 minutes before washing and incubation. It is not known for certain whether PEG treatment brings about true fusion and therefore the formation of hybrid cells on any significant scale; the work of Hazezawa *et al* (1981) suggests that while some fusion does occur, the PEG treatment is most active in promoting endocytotic uptake of the protoplasts by the much larger recipient cells. What is clear, in spite of uncertainty in this regard, is that genetic material carried by the prokaryotic element is released into the eukaryotic partner and stably expressed at high frequency (Sandri-Goldin *et al*, 1983).

Fusogenic agents other than PEG which may permit mass conjunction of cells are DEAE-dextran and inactivated Sendai virus, and these have been used variously in cell fusion technology (Poste and Nicolson, 1978) . More recently, electric field mediated fusion (Zimmermann, 1982) and laser-induced fusion (Wiegand *et al*, 1987) have been put forward as further options.

Erythrocyte Ghosts as Carriers

In this methodology the plasma membranes (ghosts) of mature, enucleate mammalian erythrocytes are partially emptied of their normal cytoplasmic constituents (largely haemoglobin) and loaded with a macromolecule of choice during hypo-osmotic haemolysis. The resultant erythrocyte carriers are then fused with appropriate recipient cells using inactivated Sendai virus or PEG so that the ghosts' contents are donated to the cells.

This technique was originally described by three different groups and was referred to variously as a means of macromolecular *injection* (Furusawa *et al*, 1974), *microinjection*

(Schlegel and Rechsteiner, 1975; 1978) and *ultramicroinjection* (Loyter *et al*, 1975), although the concept involved here must be carefully distinguished from that of mechanical microinjection. Celis *et al* (1980) have made a valuable comparison of the two kinds of microinjection, while in a later review, Celis (1984) refers to erythrocyte-ghost-mediated transfer without invoking the term microinjection at all.

In general the method has been used for the transfer of proteins or small particles (Loyter *et al*, 1975), tRNA (Schlegel and Rechsteiner, 1978) and mRNA (Boogaard and Dixon, 1983). It has been used only rarely for DNA (Iiono *et al*, 1983; Wiberg *et al*, 1983), apparently because large, unwieldy molecules like DNA cannot be loaded into erythrocytes easily by the available procedures (Celis *et al*, 1980; Celis, 1984).

The greatest importance of erythrocyte vector technology to DNA transfer techniques lies in its historical relationship to the following two methods which are based on similar principles.

Reconstituted Animal Virus Envelopes as Carriers

The idea of using inactivated Sendai virus to induce erythrocyte-cell fusion in the unloading of erythrocyte carriers has been taken even further by Loyter and colleagues (Vainstein *et al*, 1983). Using as a basis a previously reported method for reconstructing Sendai virus envelopes from constituent molecules (Hosaka and Shimizu, 1972), they have developed a means of employing these reconstituted structures as DNA carriers. Sendai virus belongs to the paramyxovirus group. Its membranous envelope bears an F glycoprotein responsible for virion-cell binding. Such an envelope therefore provides both the facility for containment and intrinsically fusogenic properties when put into contact with cells bearing the appropriate receptors.

In essence the method involves the dissolution of intact virions by the action of a detergent (Triton-X 100 or Nonidet P-40), the

removal by centrifugation of the detergent-insoluble nucleocapsid material, the reconstitution of the membrane in the presence of the DNA of interest during removal of the detergent by dialysis, and, finally, the digestion of exterior DNA by DNAse. The resultant reconstituted Sendai virus envelopes (RSVE) comprise viral phospholipids and two glycoproteins including F. For fusion to take place, it is essential that recipient cells bear sialic acid residues on cell surface glycoproteins or glycolipids. Since such sialoglycoproteins and sialoglycolipids are, in fact, present in a wide range of animal cell lines Vainstein and coworkers have dubbed the RSVE "an almost universal fusogenic syringe".

Transfer of SV40 DNA and the HSV *tk* gene has been shown to occur in a broad range of cell lines, some of which have previously demonstrated resistance to transformation by DNA-calcium phosphate complexes. Loyter's group suggest that, in theory, the RSVE method should be as efficient as intact animal viruses in the transfer of functional DNA molecules. They concede, however, that in order to reach such a goal, significant improvements would be necessary, especially in the steps involved in trapping DNA inside the envelope structures.

Liposomes as Carriers

Liposomes are laboratory-produced mono- or multilamellar vesicles in which the vesicular boundary layers are ordered assemblages of amphipathic molecules, largely phospholipid in nature, fully compatible with aqueous environments on both their internal and external surfaces and stable in water-based solutions for considerable periods. Liposomes can be manufactured by a number of methods, some of which allow efficient trapping of water-soluble high molecular weight solutes. They are therefore potentially advantageous as carriers of important biological macromolecules. Their use in this respect was first reported by Gregoriadis and Buckland (1973), who used murine cells lacking invertase and accumulating

sucrose in a sucrose-rich medium as a model for storage disease, and showed that by incubating the cells with liposomes loaded with invertase the condition could be alleviated.

Several of the properties of liposomes which are important experimental parameters in carrier technology can be varied at will: size can be adjusted from a radius of approximately 12nm for the smallest monolamellar vesicles to several μm for multilamellar types; negative or positive surface charges may be imposed by the incorporation of charged amphiphiles, and fluidity of the bilayer manipulated by alteration of its phospholipid composition. As a result, liposomes offer the worker an unusually versatile delivery system. Within the last decade, starting with the work of Mukherjee *et al* (1978) on liposome-mediated chromosome exchange, this approach has extended itself readily to specific application in purified DNA transfer.

Liposomes generally fall into one of three categories according to structure and size, factors which in turn are dependent upon the method of preparation. Multilamellar vesicles (MLV) arise when a film of phospholipid supported on the inner wall of a glass vessel is hydrated with vigorous agitation. Small unilamellar vesicles (SUV) are produced when a suspension of MLV is extensively sonicated. Neither MLV nor SUV lend themselves to DNA entrapment for two major reasons: the effective internal volume is small in both cases, and the physical rigours encountered during preparation lead to significant fragmentation of nucleic acids introduced into the aqueous phase for incorporation. On the other hand, large unilamellar and oligolamellar vesicles (LUV), which bound a relatively large aqueous volume, can be made to capture DNA very successfully if physically mild methods of manufacture are chosen from the variety available.

One method of LUV preparation which has been found applicable to DNA incorporation is an injection technique originally proposed

by Deamer and Bangham (1976) in which lipid dissolved in diethyl ether is injected into an aqueous phase warmed to a temperature at which the ether vaporises, thereby inducing liposome formation. When DNA is the solute for encapsulation, the aqueous solution is heated to 60°C and the ether solution injected very slowly and consistently at rates of 0.1 - 2.0ml/hr. The result is a population of liposomes having diameters of 0.2 - 0.5μm and a capture efficiency of 8 - 14μl DNA solution/μmole of lipid (Cudd and Nicolau, 1984).

Reverse-phase evaporation (REV), first described by Szoka and Papahadjopoulos (1978), gives rise to liposomes of similar size and holding capacity as the ether injection method, and is considered to be the most suitable and convenient preparative procedure if DNA is to be encapsulated (Straubinger and Papahadjopoulos, 1983). Aqueous buffer containing the nucleic acid of interest is mildly sonicated or extensively vortexed with a larger volume of diethyl ether containing dissolved phospholipid. This produces a stable water-in-oil emulsion in which the hydrophobic lipid tails project into the ether and the hydrophilic head groups enclose aqueous droplets. The ether is then evaporated off under reduced pressure, a process which reverses the phases, the lipids reverting to normal bilayer configuration and trapping the DNA solution previously surrounded by lipid head groups. Careful control of conditions during the formation of the original emulsion is essential if the DNA is to remain intact; the relevant parameters have been reviewed by Straubinger and Papahadjopoulos (1983) and Cudd and Nicolau (1984).

Free, untrapped DNA is often separated from the liposome fraction by gel exclusion chromatography, differential centrifugation, or flotation on discontinuous polymer gradients (Straubinger and Papahadjopoulos, 1983). In addition, liposomes of varying size may be passed through a polycarbonate filter to produce a population of uniformly sized vesicles without loss of material (Cudd and Nicolau, 1984).

The close physical association between DNA-carrying liposomes and recipient cells which is essential for delivery of the DNA occurs naturally during incubation of the two populations in physiological medium, particularly avidly if the liposomes are rich in negatively charged phospholipid (for example, phosphatidylserine) and cholesterol (Straubinger and Papahadjopoulos, 1983). Following binding, the liposomes are carried into the cell interior by endocytosis (Straubinger *et al*, 1983). It appears that liposomes do not fuse with the host cell plasma membrane unless induced to do so by the incorporation of viral fusogens into the lipid bilayer of the liposome structure or by polyethylene glycol perturbation (Struck *et al*, 1981). Based on the infectivity of liposome-carried SV40 DNA, it appears that the application of PEG (44%w/v for 90 sec) is advantageous in that it enhances effective delivery by a factor of 10, as does the addition of DMSO (25%v/v for 4 min), while glycerol (15 - 35%v/v for 4 min or more) is even more effective, enhancing delivery 100 - 200-fold (Fraley *et al*, 1981). On the other hand, one of the beauties of liposome-mediated delivery is that it can occur successfully without the agency of any such toxic fusogen.

Liposomes have been utilised for various genetic transactions. Some of the early experimental models were primarily for the purpose of demonstrating the feasibility and advantages of gene carriage and transfer by liposomes and employed only bacterial systems. One example of this approach is the work of Fraley *et al* (1979) in which the liposome encapsulation of pBR322 prior to transformation of *E. coli* was shown to make the process DNase resistant. However, there have been many reports of liposome-mediated transfections of eukaryotic cells, including plant protoplasts (Lurquin and Sheehy, 1982) and cultured murine L cells (Schaefer-Ridder *et al*, 1982). The efficiency of DNA delivery has been correlated to vesicle phospholipid composition and incubation conditions (Fraley *et al*, 1981) as well as to the type of liposome used and the stage of the cell cycle at the time of transfection (Nicolau and Sene (1982a). Nicolau and

Sene (1982b) have reviewed liposome-mediated DNA transfer and expression in eukaryotic cell systems.

Retroviruses as Carriers

Molecular cloning of functional mammalian genes in animal viruses and expression of the cloned genes in infected cells is an avenue of gene transfer which has been exploited only very recently. However, it is showing particular promise in a number of specific directions, notable progress having been made in studies of gene regulation, especially during embryonic development, as well as in studies directed towards corrective gene therapy. Consequently, viral structures have become important tools in these areas of biological research (section 1.2.3). A number of viral types have been used in these endeavours, including DNA viruses such as simian virus 40 (Mulligan *et al*, 1979; Gruss and Khoury, 1981) and vaccinia virus (Mackett *et al*, 1982). Recently, the RNA togavirus known as Sindbis virus has been suggested as a useful vector (Xiong *et al*, 1989). Most attention has been given, however, to the group of RNA viruses known as the retroviruses, and current work utilising virus particles as gene vectors is almost exclusively focused on this viral family.

The viruses which belong to the Retroviridae are unusual enveloped plus-strand RNA viruses which infect dividing cells in a range of animals including primates. The viral RNA, together with 50 - 100 molecules of reverse transcriptase, is packaged into the virus core. On cellular infection, the viral genome undergoes conversion from single-stranded RNA to double-stranded DNA, and in this form it is inserted into the genome of the host as a proviral sequence where it remains as a permanent genetic feature. There, the provirus is regularly replicated, transcribed and translated to produce viral RNA and proteins which are assembled and released as new virions, acquiring envelopes as they are budded outwards from the host cell plasma membrane. Virion production is usually at a chronic level, and

although under certain conditions retroviruses may be cytopathic, they are generally not so and the host cells are not killed; on the contrary, cellular growth may be potentiated (Bishop, 1978; Mason *et al*, 1987).

All retroviruses have a small genome (8-10kb) with a similar genetic organization. Even those which are self-contained with respect to infection and replication functions carry just a few essential structural genes, namely *gag*, *pol* and *env*, encoding the core proteins, reverse transcriptase plus associated enzymes, and envelope proteins respectively. These act in conjunction with control regions which, in the proviral genome, form long terminal repeat sequences (LTRs) flanking the structural genes on both sides (Mason *et al*, 1987; Varmus, 1988).

Many retroviruses are acutely tumorigenic as a result of the incorporation of oncogenic recombinant sequences of cellular origin into their own genomes at some stage in the past. A classically studied example is Rous Sarcoma Virus (RSV), causative of sarcomatous growths in birds, which bears the oncogene *src* in addition to the viral genes *gag*, *pol* and *env*. RSV is, however, an exceptional oncogenic retrovirus in that it is genetically complete and therefore replication competent. All others known have oncogenes substituted for viral structural gene sequences; there is consequent loss of the normal complement of *trans*-acting gene products, and the viruses are replication defective. In nature, the molecular components required for the replication and integration of such a defective virus are complemented in *trans* by a non-oncogenic replication-competent helper virus. Such a system allows a large proportion of the defective viral genome to be dispensed with, as the only *cis*-acting sequences which must be retained intact if the viral life-cycle is to continue are the LTRs, the sites for the priming of reverse transcription and a packaging signal sequence, all of which are situated in the terminal regions of the provirus. Certain naturally occurring oncogenic

retroviruses demonstrate this potential for genomic plasticity by showing extreme degrees of genetic depletion and substitution (Pimentel, 1986).

The first infective retrovirus constructs containing artificially introduced non-retroviral genes were designed and reported by Shimotohno and Temin (1981) and Wei *et al* (1981). In the latter case, a recombinant genome carrying the *cis*-acting LTR and closely adjacent regions and the *src* oncogene of Harvey murine sarcoma virus (Ha-MuSV), together with the thymidine kinase (*tk*) gene of herpes simplex virus-1 (HSV-1) linked to it in a downstream position, was constructed in bacteriophage Lambda, cloned, and used to transform mouse NIH3T3 fibroblasts by calcium phosphate transfection procedures. On superinfecting the transformed fibroblasts with a helper-independent retrovirus - Moloney murine leukaemia virus (Mo-MuLV) - new individual retrovirus were isolated. These pseudotypes were shown to be infective towards NIH3T3 cells, a single virion being capable of expressing both the *src* and *tk* genes. While the new retroviral sequence *src-tk* was non-productive (replication-defective) it could be rescued into virus particles at high titre by, again, infection with a helper-independent retrovirus. It was found to be a single RNA species 4.9 kilobases long.

In creating a fully infective but non-replicative dual gene function retrovirus of this nature, the stated purpose of Wei and coworkers was the creation of a unique means of gene transfer which would be a most useful tool in the genetic analysis of *src* and other retroviral genes, especially other oncogenes. In addition, the mode of preparation of the virus was seen as a model system for understanding how the natural and spontaneous generation of oncogenic retroviruses might have occurred during evolutionary history.

By contrast, later workers saw the powerful potential of retroviral vectors in broader terms. For example, Willis *et al* (1984) have pinpointed the general advantages of retroviral gene

delivery over transfer by other methods applicable to animals. They comment in particular on the 100% efficiency of the process of cellular entry by retroviruses. They mention also that cell types recalcitrant to calcium phosphate-mediated DNA transfer may be retrovirally infected, and that, uniquely, a single copy of the transferred genetic sequence is integrated into the recipient cell genome, simplifying studies of its organisation and expression. Miller *et al* (1984) have made similar points, drawing attention to the capacity of the retroviral structure for genomic flexibility.

The growing perception that retroviruses carrying oncogenes are naturally occurring and highly evolved eukaryote gene vectors, combined with the realisation of the extreme order of efficiency with which retroviruses are capable of infecting cells, have jointly been the key to an even more broadly based insight: that of the possible use of retroviral vectors as agents of therapeutic gene transfer to cell lines and, ultimately, to whole animals including humans (Robertson, 1986; Varmus, 1988). Natural retroviral mechanisms have now been expropriated by clinical investigators with vector designs of their own, and this aspect of retroviral development has advanced enormously in the last few years, albeit encountering unforeseen difficulties on the way. It is discussed further in section 1.2.3.

1.2.3 *APPLICATIONS OF GENE TRANSFER TECHNOLOGY*

The deliberate insertion and integration of functionally active genes into eukaryotic cells has broad-ranging and enormously important practical relevance in molecular biology, as well as in the many fields of applied science which draw upon its technology: medicine, agriculture and industry. The extensive array of transformation methodologies reviewed in section 1.2.2 and the continuous emergence of new ones in themselves attest to this. Outlined briefly below are the areas of research in which radical advances have been made that are directly attributable

to the facility for experimental gene transfer.

Gene Expression

Extremely fine analysis of the mechanisms by which gene expression is regulated has been possible as a result of the transfer of precisely engineered combinations of structural genes and control region components to recipient cells in which expression can be monitored. In fact there is almost no current biochemical research in the area of gene regulation which does not employ gene insertion techniques. This type of work has involved much use of expression vectors carrying assayable bacterial genes (section 1.2.1) fitted with various arrangements of eukaryotic regulatory elements, and measurement of their genetic activity on delivery to cultured cells (Johansen *et al*, 1984; Rosenthal, 1987). Important advances such as the identification of the roles of cis-acting promoter and enhancer DNA sequences in the regulation of transcription has come about almost entirely because of this kind of approach (Maniatis *et al*, 1987). Furthermore, headway has been made in understanding how location-specific and temporal expression is controlled by these elements, as for instance in the study by Elshourbagy *et al* (1987) of the expression of the human apolipoprotein A-IV gene according to tissue type. Just how functionally versatile such elements may be is also emerging as a key concept, as for example in the bidirectional control of a gene pair by a shared enhancer (Nickol and Felsenfeld, 1988).

Protein Structure-Function Relationships

The insertion of structural genes and gene fragments into a foreign or homologous cell system for the contrived production of the coded protein has become an important means of protein structural analysis. One of the early examples of this approach is the cell surface expression of the influenza haemagglutinin glycoprotein, an integral membrane protein, following host cell infection with a viral vector (SV40) carrying a cloned DNA copy

of the original RNA gene (Gething and Sambrook, 1981). A comparable study by Rose and Bergmann (1982) involved cell surface secretion of vesicular stomatitis virus glycoprotein. Procedures such as these have opened the way to detailed study of both viral and non-viral genes and their products which would be inaccessible to genetic analysis at their natural loci. There are many investigations in which the use of site-directed mutagenesis or chemical modification for the production of minor structural alterations, alternatively the insertion of foreign nucleotide sequences into the gene for the production of hybrid proteins, allows dissection of the molecular domains of the protein in terms of structure, function and biosynthesis. In one example, observations of intermediate filament formation in hamster lens and HeLa cells transfected with modified hamster vimentin and desmin genes has established the roles played by terminal regions of the proteins in intermediate filament assembly (van den Heuvel *et al*, 1987).

Physiological Processes

As more mechanisms of cellular physiology become understood in terms of peptide and protein activity, the use of gene transfer techniques to aid in further elucidation is increasing. One field which lends itself to this approach is that of neuropeptide processing (Thomas and Thorne, 1988).

Cellular Differentiation

Hierarchies of cellular differentiation have been elucidated by the introduction of clonal markers into primitive stem cells. In the majority of studies of this kind, the haematopoietic system has been used as a model, *neo* has been used as the gene marker, and retroviral vectors have been the means of marker insertion. The gene transfer step is accomplished *in vitro*, but only as a preliminary to reestablishment *in vivo*, this type of investigation being the first to employ gene delivery to intact animals (Cline *et al*, 1980). In a feasibility study in mice in

which foreign genetic material was introduced into pluripotent haematopoietic stem cells (the progenitors of all blood tissues), Mulligan and coworkers (Williams *et al*, 1984) pinpointed the particular usefulness of retroviral vectors for work of this kind: the stem cells to be transfected represent only 0.01% of the bone marrow population, and have not been defined morphologically, thus making a highly efficient method of gene transfer essential. Subsequent work by Mulligan's group (Dzierzak *et al*, 1988) has utilised the method to analyse lineage-specific gene expression. Dick *et al* (1985) employed this same approach to monitor cell lineages in some detail, and made clever additional use of the retroviral gene transfer mechanism: progeny from individual stem cells were unequivocally identified by their unique proviral integration site on the chromosome.

Developmental studies comparable to these have been made possible using preimplantation embryos as gene recipients, gene insertion usually being accomplished either by analogous retroviral protocols (Rubenstein *et al*, 1986) or microinjection (Gordon and Ruddle, 1983). In an ingenious variation, Bosselman *et al* (1989) have combined the two in order to overcome difficulties previously encountered in introducing genes into early embryonic chickens: they microinjected nonreplicative retroviral vectors into stem cells in the unincubated embryo, with a considerable degree of success.

Role of Oncogenes in Carcinogenesis

Twenty oncogenes have been identified in oncogenic retroviruses (Duesberg, 1987). In addition, cellular protooncogenes that have been subjected to somatic mutation, amplification or rearrangement, thereby creating potential oncogenes, have been found in a variety of tumours (Bishop, 1987). Using cell culture as a model system, and phenotypic changes such as cellular morphology and growth in soft agarose as indicators of neoplasticity, oncogenes and putative oncogenes have been tested

for their ability to transform primary cell cultures, usually of embryonic rodents. The sequences to be tested are most often inserted into the cells via plasmid constructs, delivered by means of simple methods of transfection such as the calcium phosphate-DNA coprecipitate protocol. As a result of this approach, many oncogenes have been identified, a huge body of information on oncogene expression has been accumulated, and important principles derived concerning the way in which tumorigenesis is genetically determined. For instance, it is now known that although some oncogenes may unilaterally bring about cancerous changes (Hjelle *et al*, 1988), many others do so only in cooperation with a second or third oncogene (Land *et al*, 1983). From a similar experimental base of molecular genetics it is also now clear that the products of oncogenes and protooncogenes are most frequently growth factors or growth factor receptors (Guroff, 1987) and the insight this unifying concept allows into the possible mechanism of carconogenesis suggests structured ways of further investigating the problem.

Therapeutic Gene Replacement

Within the arena of gene transfer, the drive to devise and perfect ways of delivering corrective genes to genetically defective animals has been perhaps the most persistent. Its fundamental operation depends upon the dual availability of the facility for gene delivery to whole organisms and efficient stable integration. It has therefore relied upon one of three approaches: (i) microinjection or retroviral delivery of genes to the germ line, (ii) retroviral *in vitro* transformation of somatic cells subsequently transplanted into a recipient animal, and (iii) direct *in vivo* vehicle-mediated transformation of somatic tissue via the circulatory system. A possible advantage of (i) is that it results in permanent, vertically transmitted therapeutic change. It is typified by the work of groups led by Jaenisch and Costantini respectively (Jaenisch, 1988; Constantini *et al*, 1986). Both (ii) and (iii) effect change which is restricted to the recipient organism, but the potential

for experimental control and tissue targeting in these procedures appears to be distinctly advantageous, and the risks and costs incurred would be less than for (i), thus making them more likely candidates for use in humans.

Not surprisingly, the haematopoietic system has been envisaged as a convenient manipulable system for the development of somatic therapy involving retroviral transformation of transplantable tissue, and it has been suggested and accomplished using bone marrow by a number of laboratories led respectively by Mulligan, Miller, Anderson and others who were active in establishing the necessary technology for studies of blood cell lineages, mentioned above. Working with the bone marrow system, Hock and Miller (1986) made an important contribution to retroviral therapy in pioneering the design of "suicide" infections, using packaging-defective helper viruses in conjunction with the replication-defective retroviral vector, thereby achieving transformed cells completely lacking any residual contaminating viruses. More recently, the same group of workers has turned attention to the prospects for therapeutic implantation or grafting of genetically engineered skin fibroblasts into recipient patients; they have initiated investigations by correcting specific medically important gene deficiencies in human fibroblast cultures (Anson *et al*, 1987; Palmer *et al*, 1987), while other workers have undertaken fibroblast therapy in murine whole animal models (Selden *et al*, 1987). The facility these approaches offer for characterisation of the transformed cell line before transplantation is particularly appealing.

Gene therapy via the circulatory system is of continuing interest, but less well developed than cell transplantation techniques, especially with respect to whole animal trials. Of the gene delivery methods developed to date the only ones which lend themselves to circulatory use are those involving vehicle-mediated transfer, and the suggested approach closely resembles that of vehicle-mediated drug delivery which has been

investigated in some depth over a number of years (Gregoriadis, 1977; 1984). Many studies have been devoted to the potential these delivery systems have for tissue-specific targeting via modifications of the vehicular membrane, particularly the inclusion of appropriate antibodies into liposomes (Gregoriadis and Meehan, 1981; Peeters *et al*, 1988), but few have demonstrated targeted DNA delivery by vehicular means. One which has, but only with a cell culture model, is the work of Gitman *et al* (1985) which showed the targeted delivery of SV40 DNA, via Sendai virus envelopes incorporating insulin, to virus receptor-depleted hepatocytes. Other studies have dwelt upon the potential of certain tissues to be used as targets. For example, Smirnov *et al* (1986) have suggested that denuded areas of blood vessel walls might make a useful target for liposome and erythrocyte carriers. Using similar thinking but a viral vector approach, Anderson's laboratory is currently looking at the transducibility of endothelial cells, on the basis that the immediate contact the endothelium has with the circulating blood makes it an attractive *in vivo* target tissue for retrovirally delivered genes (Zwiebel *et al*, 1989).

A distinction must be made between extracellular targeting of genes to specific cells or tissues and intracellular targeting of genes to homologous sites within the chromosome. The latter is crucially important if precise gene correction is to be achieved, yet most methods of gene insertion used at the present time lead to uncontrolled and illegitimate recombination events in mammalian systems (Kato *et al*, 1986), while retroviral gene transfer is a system of naturally controlled random insertion (Dick *et al*, 1985). Moreover, there is evidence that although DNA may be targeted to sites of homology at high frequency under certain conditions (Thomas *et al*, 1986), the result may be the production of mutation in the cognate gene (Thomas and Capecchi, 1986). Targeted recombination is therefore an area of great importance within the field of gene therapy; it continues to produce interesting results from systems in which attempts are made to represent the genetic environment more normally than

previously (Adair *et al*, 1989), and may lead to more prolific investigations in the near future.

Transgenic Animals and Plants

The introduction of foreign genes into the germ line and successful expression of the transferred gene has led to genetic manipulation of animals on a major scale. In a review of the subject, Jaenisch (1988) gives details of the techniques employed. They can be summarised broadly as follows: (i) injection of recombinant DNA into the pronucleus of a fertilised egg; (ii) retroviral gene transfer to pre- and post-implantation embryos, leading to some germline transformation, and (iii) gene transfer by a variety of methods into embryonic stem cells established *in vitro* from explanted blastocysts, followed by reintroduction of the transformed cells into a host blastocyst for colonisation and contribution to the germ line. "Transgenic technology" (Jaenisch, 1988) clearly provides a model system for many areas of study, contributing to advances being made in areas of endeavour already discussed above. In addition, however, the real possibility exists of generating transgenic animals of commercial importance. Extension of the technology from the common murine model (Palmiter *et al*, 1982) to rabbits, sheep and pigs has been accomplished (Hammer *et al*, 1985), and it seems likely that such strategies as the harvesting of medically important proteins from the milk of transgenic cows might well be implemented in the near future, as has been shown feasible and successful in mice (Simons *et al*, 1987).

In the plant world, transgenic individuals are rather more easy to produce than in the animal kingdom, a situation which is leading to a massive increase in the scope of plant molecular biology, as well as commercial investment in it. Transgenic plants have been produced by DNA injection into floral tillers (De la Pena *et al*, 1986), but more commonly by simple transfection of somatic protoplasts, cells and tissues subsequently allowed to develop into new plantlets via the

production of callus and differentiated tissue in culture (Jones *et al*, 1985; Rogers *et al*, 1986). As with transgenic animals, many fundamental genetic studies have made progress as a result of the technology (Willmitzer, 1988); in addition, transgenic plants are now being designed for horticultural and agricultural use (De Block *et al*, 1987).

1.2.4 *THE NEED FOR NEW APPROACHES*

The plethora of methodologies available for the introduction of genes into cells, tissues and whole organisms does not yet mean that a suitable technique exists for every possible purpose. Simple transformation techniques for genetic studies of cells *in vitro* are well developed, and perhaps require little augmentation. However, the complex requirements of *in vivo* transformation are only just being realised, and are certainly not met by any one approach. This is amply illustrated by the fact that human gene therapy is still thwarted by biological obstacles, not merely slowed down by the difficulties being encountered in attempting to resolve the ethical problems involved. Table 1.1 summarises the important limitations besetting each of the major gene transfer methods at the present time. It shows that while techniques for the delivery of naked DNA are the ones most useful for cell culture work, they are not suitable for application to whole animal systems since they involve protocols which are inappropriate or too aggressive; the only gene transfer methods which mimic natural conditions sufficiently closely to be useful *in vivo* are those involving vehicle-mediation, of which the retroviral mechanism alone is efficient enough for serious practical consideration.

The work described in this thesis is concerned with the development of a new approach to gene delivery, that of transfer of soluble naked DNA across the plasma membrane by receptor-mediated endocytosis. The work presented constitutes a study of the feasibility of this approach, modelled *in vitro*, but its

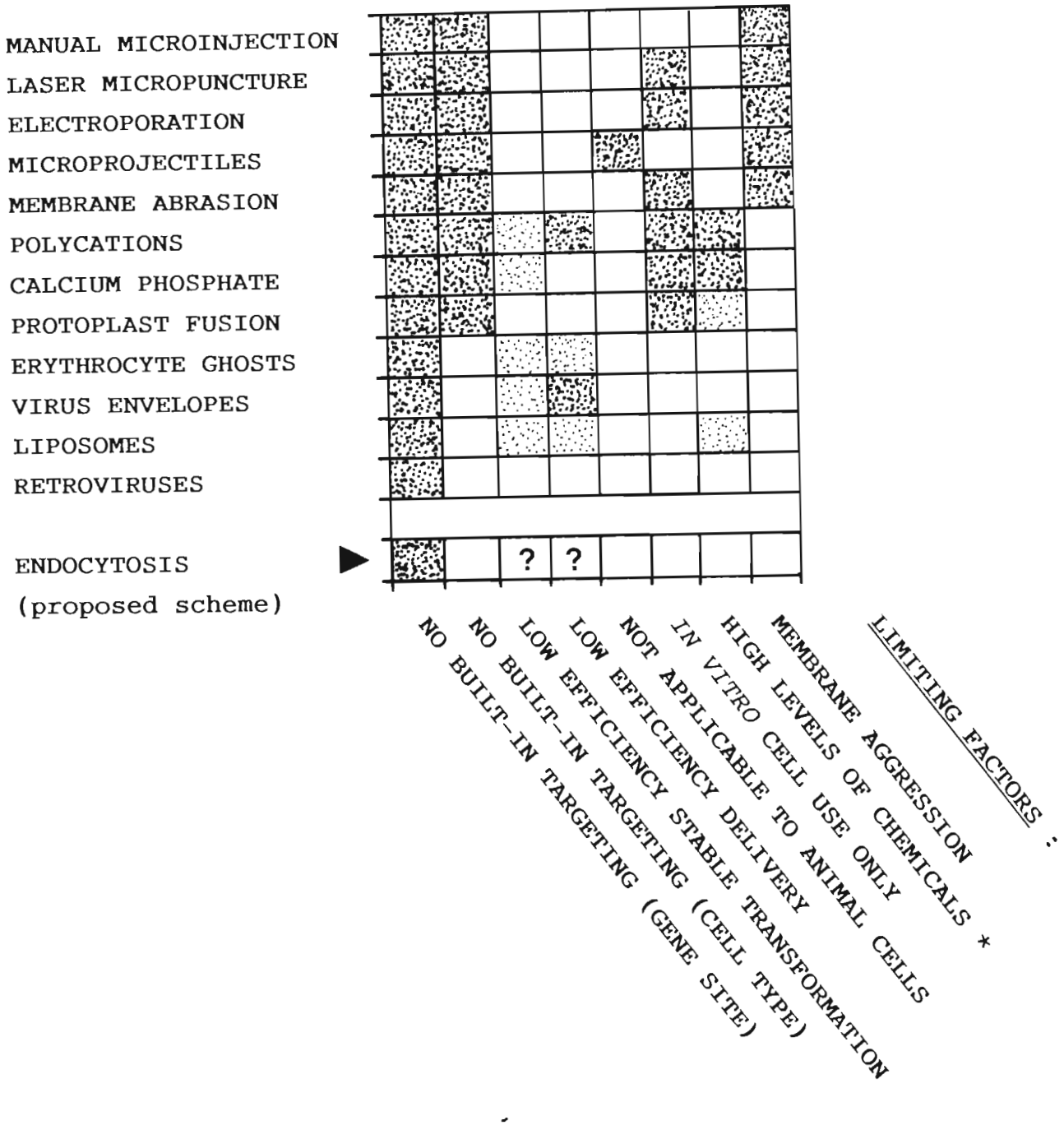


Table 1.1 Negative factors restricting applicability of current gene transfer methods

* Perturbations of gene expression have been documented for the calcium phosphate procedure (Kerbel *et al*, 1987; Pine *et al*, 1988) (section 1.2.2).

[▒] = Serious limitation to applicability of method
[▫] = moderate or partial limitation

potential does lie beyond cell culture systems; the possibility exists for its use in the targeted delivery of genes to mammalian cells *in vivo* via the circulatory system, without the involvement of vehicle-mediation.

1.3 GENE TRANSFER BY RECEPTOR-MEDIATED ENDOCYTOSIS

1.3.1 *CONCEPTUAL FOUNDATION OF THE THESIS*

This thesis is based upon the idea that the efficiency and specificity of internalisation afforded by the process of receptor-mediated endocytosis (Goldstein *et al*, 1979) might be exploited for the targeted delivery of genes to mammalian cells in a directed manner. Implicit in the idea is the suggestion that if DNA in the naked state could be bound to a soluble protein ligand recognisable by its cognate receptor at the plasma membrane surface, it would be carried into the cell interior. This concept has been put forward previously by Pastan and coworkers (Cheng *et al*, 1983) in a communication reporting the development of a method for covalent attachment of DNA to protein. By contrast, non-covalent reversible binding of DNA to protein was envisaged here in the belief that it might offer advantages in the cytoplasm in the event of successful transfection.

At the outset of the work it was known that the complexing of DNA to protein in a non-covalent manner is possible if the protein is first modified to a condition of extreme basicity by reaction with certain water-soluble carbodiimides. During the course of earlier investigations using carbodiimides to couple proteins to various small molecules, it had been observed that control proteins treated with carbodiimide in the absence of a second molecule gave rise to products capable of binding negatively charged materials including DNA. This type of interaction is explained by the addition of positively charged

N-acylurea carbodiimide moieties at side chain carboxyl groups of aspartate and glutamate residues (Timkovitch, 1972), and the formation of tenaceous salt bridges between these groups and ionised phosphodiester backbone regions of DNA.

The proposed approach thus involved the preparation of a modified protein which could act as both receptor specific ligand and DNA carrier. It was modelled using insulin as the protein of choice. On account of its small size, limited number of modification sites and crucial conformation-dependent receptor binding role, it was thought that insulin itself should not be carbodiimide treated; instead it should be cross-linked to bovine serum albumin (BSA) which has a capacity for extensive carbodiimide modification (Table 1.2) and could be converted to its *N*-acylurea derivative prior to coupling. In this way a conjugated protein would be created in which the receptor binding and DNA binding functions were relegated to separate surfaces on one macromolecule. The conceptual design of this conjugate is illustrated in Figure 1.1.

Table 1.2 Relative potentials of insulin and albumin for carbodiimide modification

PROTEIN/ PEPTIDE	SOURCE	MOLECULAR WEIGHT	TOTAL NUMBER OF Asp & Glu RESIDUES
INSULIN (monomer)	pig, cow human	5700	4 *
ALBUMIN	cow	68000	97 **

* (Smith, 1972)
** (Brown, 1977)

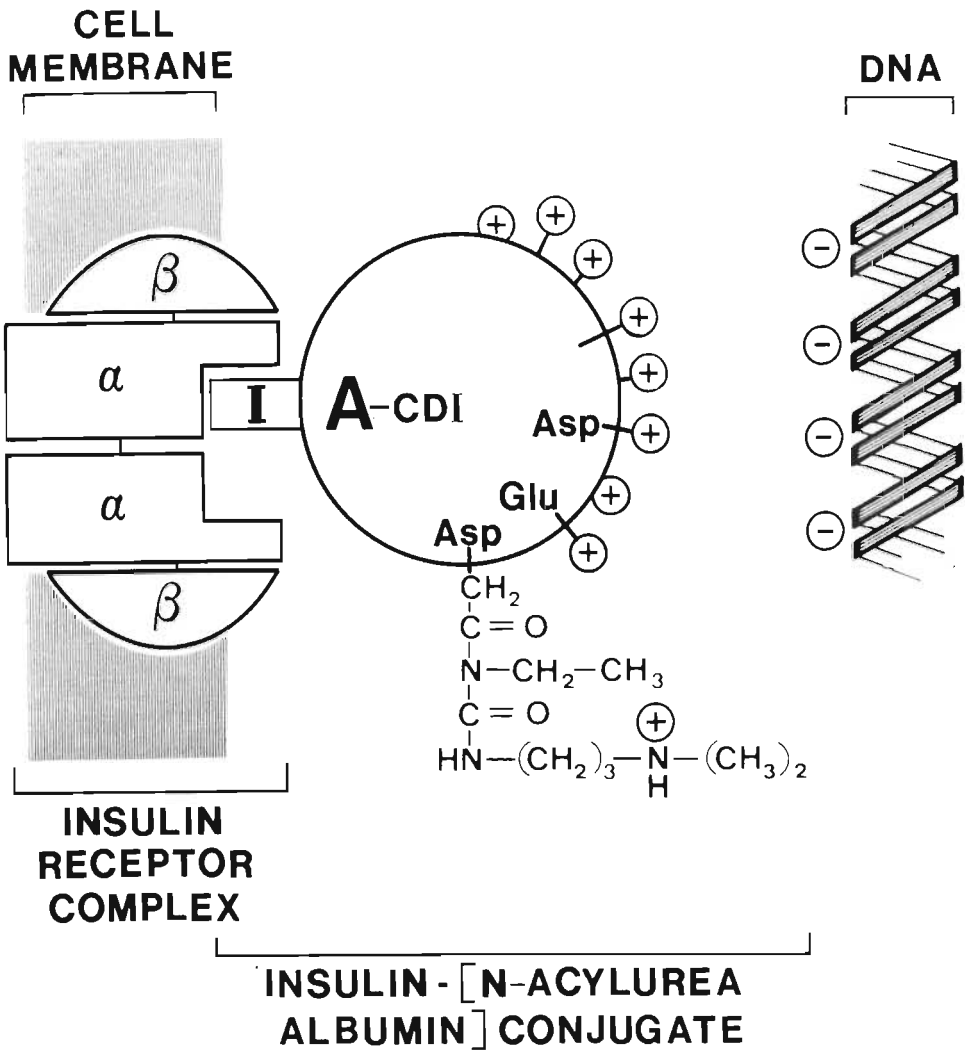


Figure 1.1 Binding interactions of a conjugate consisting of a small polypeptide ligand such as insulin covalently cross-linked to *N*-acylurea albumin

The ligand component retains its facility for specific receptor binding. The *N*-acylurea albumin component, produced by carbodiimide modification of serum albumin, bears numerous positively charged *N*-acylurea moieties (⊕) in place of acidic amino acid side chain carboxyl groups (aspartate and glutamate residues) and readily interacts electrostatically with negatively charged DNA phosphodiester backbone (⊖).
I, insulin; A-CDI, *N*-acylurea albumin; α and ß, insulin receptor subunits.

1.3.2 *SERUM ALBUMIN*

Serum albumin is one of the major proteins of the vertebrate circulatory system, maintaining continuous contact with extracellular surfaces. It is a resilient, multifunctional protein which provides buffering and protection to other, more sensitive, molecular components of the organism. In addition, it plays more specific roles such as the solubilisation of long chain fatty acids, the binding of bilirubin and toxins, and ultimately the provision of amino acids to peripheral tissues by its own degradation.

The primary structure of serum albumin was established by Brown and coworkers (Brown, 1978); its conformation and active sites have been investigated by Peters and colleagues (Peters and Reed, 1978). The bovine serum albumin molecule is composed of 582 amino acid residues and has a molecular weight of 68500; other serum albumins show slight variations. Serum albumins are prolate ellipsoids in shape, 14.1 X 4.2nm in size, consisting of three domains in linear sequence. The domains are very similar to one another, and are essentially a triply repeated series of loop structures which have arisen during evolution by gene duplication (McLachlan and Walker, 1977). The domains have developed differences of hydrophobicity and net charge, and ligand binding is domain specific. In human serum albumin, the first domain binds indole compounds and mixed disulphides, the second domain binds bilirubin, and the third domain binds long chain fatty acids.

Serum albumin is very soluble in aqueous solvents and shows exceptional stability in solution, resisting extreme conditions which would denature most other proteins. For instance, it is particularly heat tolerant, withstands the action of agents like 6M urea which tend to disrupt hydrogen bonds, and requires a concentration of trichloroacetic acid (TCA) as high as 30%w/v for precipitation. Its ruggedness and stability, combined with the fact that it is readily available and inexpensive, make it a

popular choice when a model protein is sought for laboratory procedures. Furthermore, it is an important reference standard in protein assays because it is one of the few common polypeptides free of carbohydrate. It is widely used to protect low concentrations of other molecules (for example, enzymes, antigens and antibodies) from denaturation, and to prevent their adsorption to glass vessels. It is a convenient blocking agent in Western-type blotting techniques.

Serum albumins have been put to a number of specialised uses. For example, they have found widespread application as carriers for haptenic molecules in the preparation of antibodies. A variety of proteins has been employed for this purpose, but serum albumins and keyhole limpet haemocyanin (KLH) remain the ones of choice. A major reason is that conjugates prepared with human, bovine and rabbit serum albumins are far more soluble than those prepared with other inexpensive proteins such as γ-globulin and egg albumin; the latter frequently precipitate out during synthesis (Erlanger, 1980). Serum albumins have also become much used as carrier molecules for drugs, the pharmacological potential of albumin's humoral role adding further advantage to the ease with which conjugated derivatives can be prepared. Current methodology includes the massive polymerisation of albumin into microspheres from which drugs may be released gradually and continuously (Tomlinson and Burger, 1985).

In the work described in this thesis, bovine serum albumin was chosen to act as a dual purpose carrier in accordance with the model proposed in Figure 1.1. Carbodiimide modification of the albumin to produce an *N*-acylurea albumin was undertaken in order to render the protein capable of DNA binding and carriage, while its covalent conjugation to insulin was undertaken so as to facilitate targeting and insertion of the complex into cells bearing insulin receptors.

Throughout the work, references to *albumin* signify bovine serum

albumin (BSA).

1.3.3 *INSULIN AND THE INSULIN RECEPTOR*

In most vertebrate species insulin is a major anabolic polypeptide hormone, secreted by the ß-cells of the pancreas, which is capable of eliciting a remarkable array of biological responses. It is an attractive choice of proteinaceous molecule for use as a model in the work described here because it binds to specific plasma membrane receptors on a wide range of target cells, following which the endocytotic uptake of the insulin-receptor complex and its temporary sequestration in the endosomal apparatus occurs rapidly and inevitably (Bergeron *et al*, 1985; Sonne, 1988). The insulin receptors are abundantly embodied in the membranes of liver, fat and muscle cells, and significantly represented in a number of other tissues such as blood and brain (Cahill, 1971); there are few cells which lack them completely (Rosen, 1987).

Insulin was the first functional protein to be sequenced, a major analysis by Sanger and coworkers (Ryle *et al*, 1955) now considered to be classic. Insulin is composed of 51 amino acids residues in two peptide chains, A and B, linked by two disulphide bridges and containing a third intrachain bridge. It has a molecular weight of 5700. Its full primary structure is given in Figure 1.2.

Crystallisation studies by Schlichtkrull (1956a; 1956b) established that the presence of zinc was required for insulin crystal formation: two zinc molecules for every six molecules of insulin. Meticulous X-ray crystallographic analysis by Hodgkin and her colleagues subsequently led to the elucidation of the tertiary structure of the zinc-insulin complex (Blundell *et al*, 1971; 1972). It was shown to be a hexameric configuration comprising a trimer of dimers in which the zinc is required to coordinate the three dimers but takes no part in dimer formation

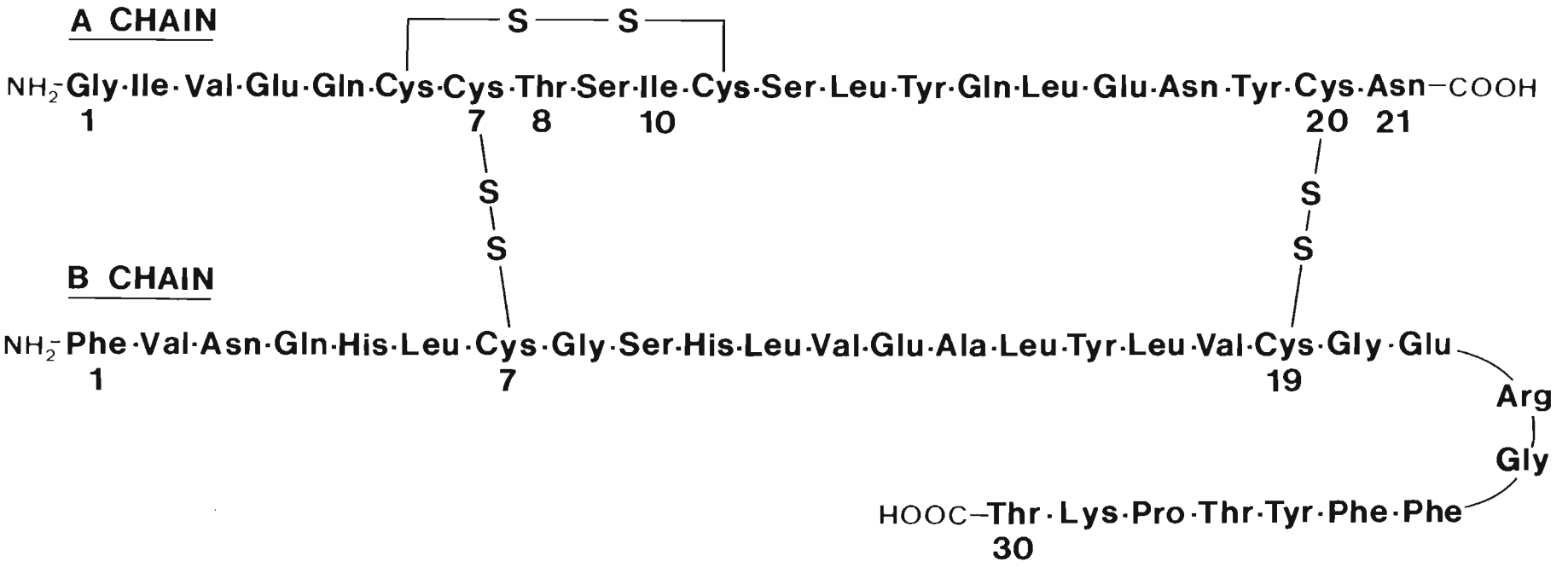


Figure 1.2 The primary structure of insulin (Ryle *et al*, 1955).

The general structure shown including the disulphide linkages between cysteine residues is applicable to a wide range of insulins.
The specific amino acid sequence shown is for human insulin.
Variation in porcine insulin: B30 = Ala (Smith, 1972).
Variation in bovine insulin: A8 = Ala; A10 = Val; B30 = Ala (Smith, 1972).

Figure overleaf/...

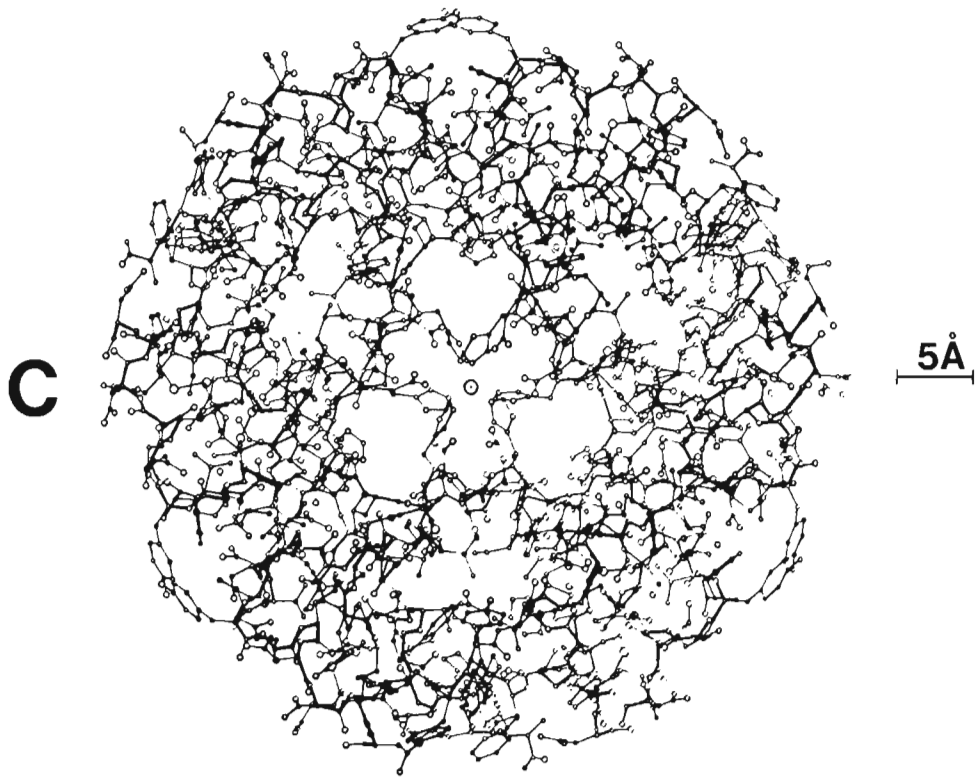


Figure 1.3 The tertiary structure of insulin (from Blundell *et al*, 1971; 1972).

The rhombohedral insulin crystal is a hexamer composed of three equivalent dimers related to each other by 3-fold rotation. The dimer itself constitutes an asymmetric unit. Perpendicular to the 3-fold crystal axis and intersecting it, there is a 2-fold axis of symmetry which relates the two monomers of the dimer and neighbouring dimers of the hexamer.

A. Projection of the complete insulin monomer, including side chains, along the 3-fold axis which is perpendicular to the plane of the paper. The 2-fold axes lie in this plane and are labelled OP and OQ.

B. The insulin dimer viewed down the 3-fold axis.

C. The complete insulin hexamer, which constitutes the rhombohedral unit cell. There are 2 zinc ions situated 17 - 18Å apart on the 3-fold axis (⊙) and related by the 2-fold axes. Each zinc ion has bound to it an imidazole nitrogen of one of the B10 histidines of each of the dimers; thus each zinc has bound to it 3 imidazole rings and is coordinated to all 3 dimers.

Note: the view of the insulin molecule given in Figure 1.4 is in a direction perpendicular to the 3-fold axis.

itself (Figure 1.3). The three-dimensional structure of monomeric insulin, which does not crystallise, is inferred from that of the hexamer. It is a well defined, relatively inflexible globular structure, having a buried core of hydrophobic residues and a surface characterised by both polar and non-polar regions. The non-polar regions comprise two patches of clustered hydrophobic residues, one involved in dimer formation, the other in dimer-dimer interactions in the hexamer. The biological significance of the different forms of insulin has been discussed by Blundell *et al* (1972). While circulating humoral insulin is in dilute solution and considered to be monomeric, insulin stored in the pancreas is granulated, with a crystal structure showing a cubic or rhombohedral close packing of zinc insulin hexamers that closely resembles the insulin crystals produced *in vitro*. The association of monomers in storage is considered to decrease the rate of enzymatic proteolysis, increase thermodynamic stability and provide a convenient concentrated store of the hormone.

Knowledge of primary and tertiary structural features, comparisons of amino acid sequence and biological activity data from insulins of different species, and activity measurements using chemical analogues, have all combined to provide a basis for interpreting insulin structure in functional terms. The result has been the definition of a biologically active region on the molecule consisting of a cluster of invariant residues where receptor binding and other reaction events occur. The numerous studies contributing to this structure-function analysis have been reviewed in detail by Gammeltoft (1984), who has pointed out two important assumptions implicit in the work. Firstly, that the form of insulin actively binding to the receptor is monomeric, substantiated by the observation that the physiological concentration of insulin is 10^{-11}-10^{-9}M while the K_d of dimerisation is in the region of 10^{-6}M. Secondly, that the monomer in solution exhibits the structure inferred from the crystal; this assumption is supported by studies of the chemical reactivity (Blundell *et al*, 1971) and circular dichroism

(Blundell and Wood, 1982) of insulin in solution.
The deduced receptor-binding site (Figure 1.4) includes the mainly hydrophilic A chain residues A1 (Gly), A5 (Gln), A19 (Tyr) and A21 (Asn) and adjacent hydrophobic B chain residues B12 (Val), B16 (Tyr), B24 (Phe), B25 (Phe) and B26 (Tyr). The same non-polar B chain residues are also involved in dimer formation (Blundell *et al*, 1972), suggesting that receptor binding might be analogous to dimerisation, and that the loss of non-polar surface residues in the formation of the complex is likely to be an important driving force of the reaction. The binding forces involved in both types of association have been shown to include hydrogen bonding in a ß-pleated sheet manner between main polypeptide chains and hydrophobic interactions of side chain groups (Pullen *et al*, 1976). The thermodynamic behaviour of insulin-receptor interaction (Waelbroeck *et al*, 1979) supports Blundell's hydrophobicity concept, but indicates that additional binding mechanisms must exist in the insulin-receptor complex: the K_d for receptor binding (10^{-9}M) is significantly greater than the K_d for dimer formation (10^{-6}M). In this regard, Gammeltoft (1984) has suggested the possibility of ionic bonding between charged residues in the A chain of the insulin binding site and residues at the receptor binding site, and Waelbroeck (1982) has provided evidence for a contribution to binding by the terminal amino group on the A chain (residue A1). Furthermore, the recent finding of Derewenda *et al* (1989) that phenol binds and stabilises helix at the B1-B8 position in the insulin hexamer raises the possibility that a tyrosyl side chain from the insulin receptor interacts with the same B chain N-terminal region.

That insulin acts at the level of the plasma membrane had been suggested by Langley in 1905, but receptors were not identified for a further sixty-five years, when a spate of work involving receptor-binding assays with radioiodinated insulin was initiated by investigators such as Cuatrecasas (1971) and Freychet *et al* (1971). Assays were founded on the principle that receptor binding of physiological levels of [^{125}I]-insulin

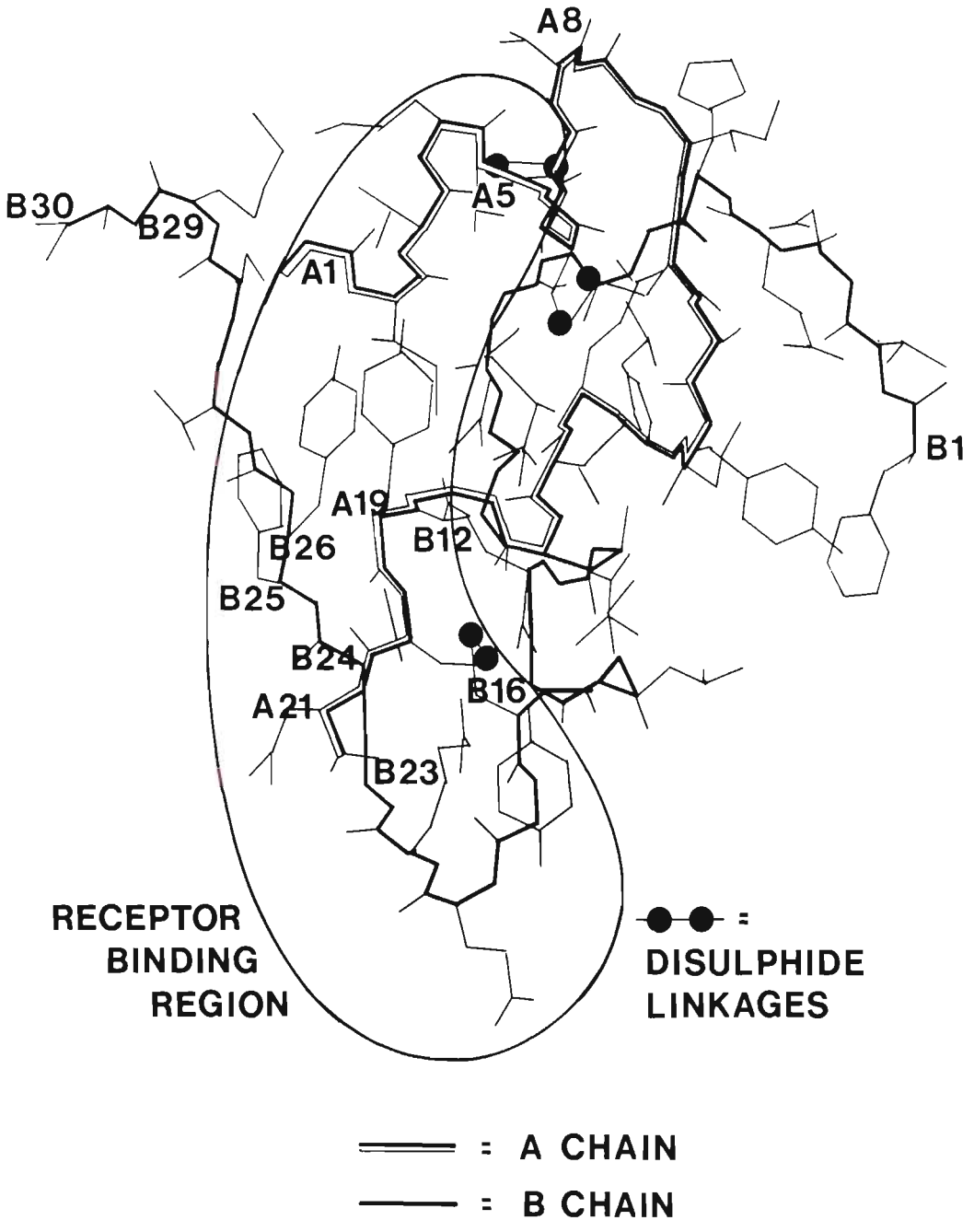


Figure 1.4 Biologically active amino acid residues of the insulin molecule (after Gammeltoft, 1984).

The outlined surface area is involved in receptor binding and includes $A1^{Gly}$, $A5^{Gln}$, $A19^{Tyr}$, $A21^{Asn}$, $B12^{Val}$, $B16^{Tyr}$, $B23^{Gly}$, $B24^{Phe}$, $B25^{Phe}$, and $B26^{Tyr}$.

Residues $B29^{Lys}$ and $A8^{Thr}$ ($A8^{His}$ in chicken and fish insulins) lie just outside the receptor binding region but may have relevance to receptor binding kinetics (Pullen *et al*, 1976; Cutfield *et al*, 1979) as discussed in section 4.4.

can be inhibited by concentrations of insulin or its derivatives 10^2 - 10^5 times higher, whereas it cannot be inhibited in this way by other peptide hormones. Interpretation of assay results laid emphasis on the dual concept of binding specificity and receptor saturability (Roth, 1973). Most assays demonstrating the presence of receptors have been conducted on isolated cells, membrane fragments or solubilised receptors *in vitro*, although some whole tissues have been used, liver for example (Terris and Steiner, 1976), and even whole organisms (Zeleznik and Roth, 1977). Initially, receptors were defined by their functional and kinetic characteristics; physical structure remained obscure. In the later 1970s, however, receptors were identified by both light and electron microscopy using insulin labelled in various ways for visual detection; for instance with ferritin (Jarret and Smith, 1974), rhodamine (Schlessinger *et al*, 1978, and ^{125}I (Bergeron *et al*, 1979). At the same time, a great deal of information about molecular structure began to be revealed by means of new biochemical approaches involving methods such as affinity labelling, affinity chromatography and specific immunoprecipitation.

The structural concept of the insulin receptor from which a current model has been derived (Pessin *et al*, 1985) was independently put forward by Cuatrecasas' research group (Jacobs *et al*, 1980) and Czech and coworkers (Massague *et al*, 1980) on the basis of a variety of biochemical evidence. Confirmation of the proposed structure has come since from the cloning and sequencing of cDNA for the human insulin proreceptor (Ullrich *et al*,1985; Ebina *et al*, 1985). The model describes a heterotetrameric disulphide-linked integral membrane glycoprotein that contains subunits of two types, α and ß, giving the structure $(\alpha\beta)_2$, and suggests that this represents the minimum covalent receptor unit. The two α-subunits (M_r 125000-135000) are linked together centrally, each flanked by and linked to a ß subunit (M_r 90000-98000). The α-subunit carries the insulin recognition site; the ß-subunit includes a tyrosine kinase function and a stretch of 23-26 highly

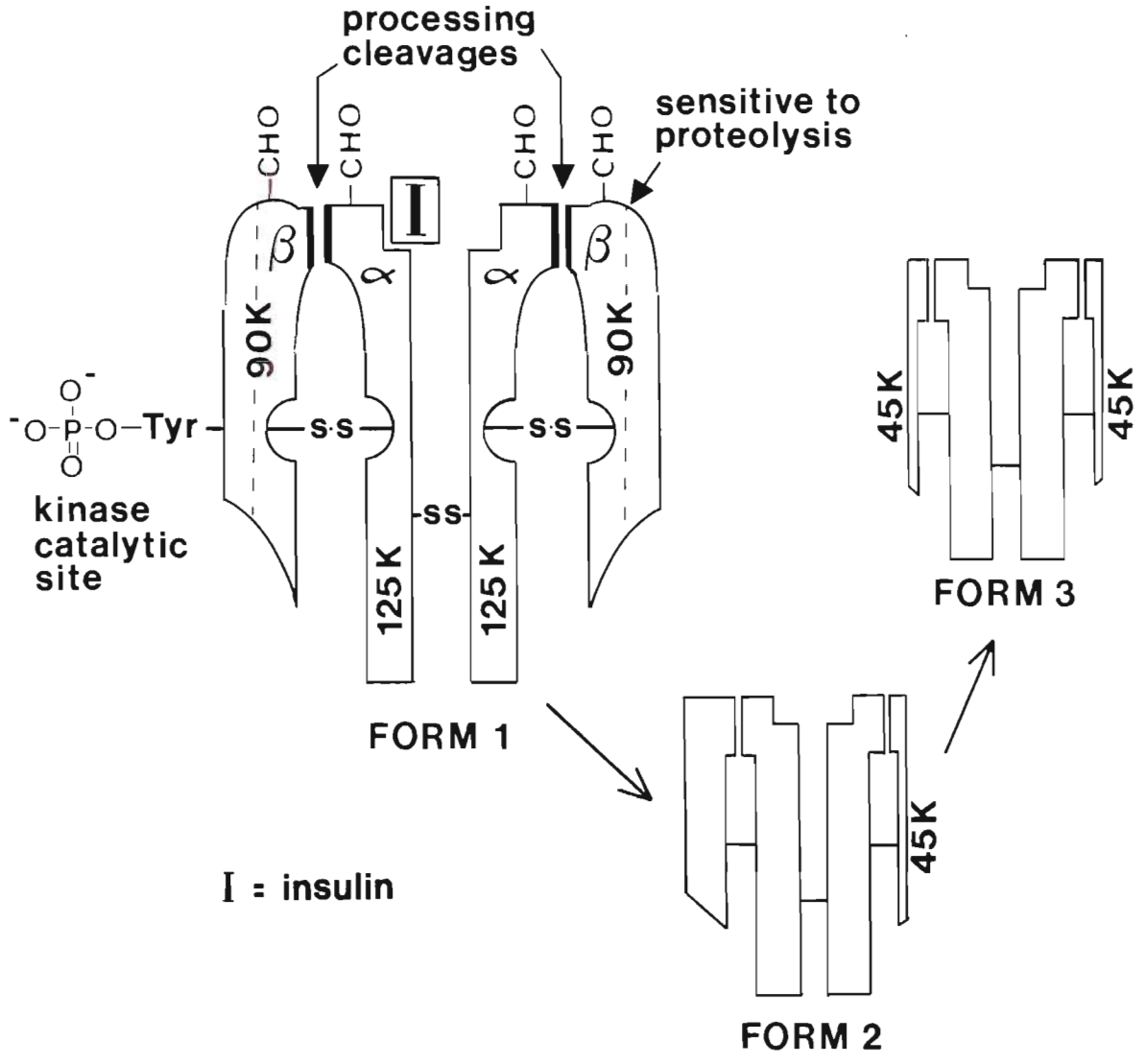


Figure 1.5 Minimum covalent subunit structure of the native insulin receptor and two major proteolytically cleaved forms (model proposed by Pessin *et al*, 1985).

Form 1 represents the hypothetical structure of the native insulin receptor, which consists of α and ß subunits linked by disulphide bonds into a heterotetrameric complex. The central disulphide bond linking the two symmetrical halves of the receptor is extremely sensitive to reductant. Adjacent α and ß subunits on each flank of the receptor complex are believed to the result of cleavage of one continuous polypeptide chain during processing. The ß subunit is phosphorylated upon insulin activation in intact cells or detergent extracts and appears to contain one or more tyrosine kinase catalytic sites. The ß subunit contains a region at about the centre of its amino acid sequence which is extremely sensitive to elastase-like proteases. Cleavage of one ß subunit yields Form 2, while cleavage of both ß subunits yields Form 3. Forms 2 and 3 are often major constituents of the total receptor population in isolated membrane fractions.

hydrophobic amino acids representing a single transmembrane domain that anchors the entire receptor in the phospholipid bilayer at the cell surface. The model is depicted in Figure 1.5.

In studies of insulin activity, the presumption is generally made that all the biological responses mediated by the hormone are initiated in the receptor, and therefore that it is the structure and function of this macromolecular complex which holds the key to cellular signalling mechanisms. In fact it is now accepted that the receptor is a multifunctional protein involved in hormone binding, transfer of information across the plasma membrane, internalisation of hormone for subsequent degradation, autophosphorylation as well as phosphorylation of cellular substrates, and possibly also delivery of message from one location to another within the cell interior. The many areas of investigation devoted to the elucidation of the receptor machinery and its cascade of physiological effects have been surveyed recently (Rosen, 1987) but are not broadly reviewed here. However, the sequel to insulin-receptor recognition which is most crucial to the DNA transfer mechanism proposed in this thesis - that of endocytotic uptake - is enlarged upon in the following section.

1.3.4 *RECEPTOR-MEDIATED ENDOCYTOSIS OF INSULIN*

Insulin is only one of many molecules that have been observed to enter cells by means of receptor-mediated endocytosis. Such molecules fall into a number of categories and include substances extrinsic to the natural life of the cell such as toxins and viruses in addition to intrinsic biomolecules like hormones, growth factors and transport proteins. There are a number of pathways of receptor-mediated endocytosis, but the model against which others are usually compared is that of low-density lipoprotein (LDL) entry into cells; postulated, verified and elaborated by Goldstein, Anderson and Brown in the first

clear delineation of the process (Goldstein *et al*, 1979). Differences between specific ligand-directed pathways are found both on the cell surface and in the vesicular compartments in which the ligands are variously delivered to intracellular sites, recycled to the cell membrane, or degraded.

A broad understanding of endocytotic mechanisms was developed during the early 1980s, coincident with the emergence of molecular models of receptor structure and function. The two areas of investigation were clearly mutually reinforcing. Pastan and Willingham (1985) have reviewed the process of receptor-mediated endocytosis in general terms, and a concise synthesis is given here. The endo- and exocytotic paradigm is outlined in Figure 1.6A. Most cultured cells have a distinctive range of receptors in their plasma membranes. Some types of receptor (for example, the LDL receptor) appear to exist only clustered together in clathrin-coated pits (Figure 1.6B and 1.6C); others are distributed randomly across the membrane surface, moving about the membrane and through coated pit sites by diffusion. Receptors of this latter type migrate with a diffusion constant of 2-8 X 10^{-10} cm^2/sec and pass through a coated pit approximately every 3 seconds. In the majority of cells in culture, 1% of the membrane surface is occupied by coated pits, 500-1000 pits per cell, each coated pit giving rise to an intracellular vesicle via membrane invagination every twenty seconds or so in a constitutive manner, that is, at a rate which is independent of receptor occupancy. Since ligands bind to their receptors at high affinity, and receptors are endocytosed in such an efficient manner, ligand uptake is extemely rapid. Less than a minute after a ligand such as a growth factor or hormone is added to a cell, it can be found in an intracellular vesicle. The vesicles are called *receptosomes* (Willingham and Pastan, 1980) or *endosomes* (Marsh and Helenius, 1980). They are biochemically and morphologically distinct from lysosomes, Golgi, or other organelles, and act as agents of transport specifically for ligands and their receptors. When first formed they are 200nm in diameter and their internal pH

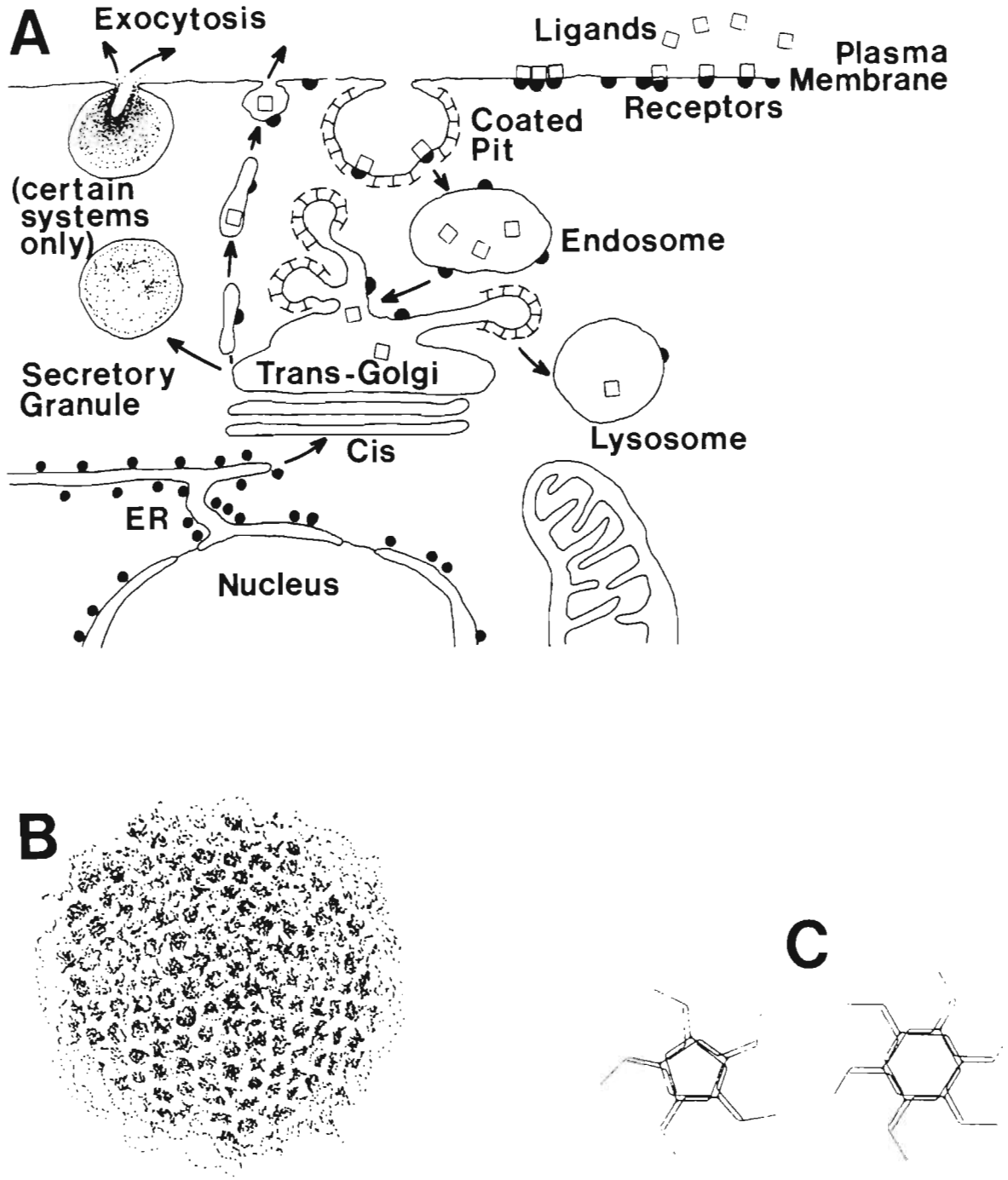


Figure 1.6 Receptor-mediated endocytosis

A. Diagrammatic summary of the classic morphological pathways of endocytosis and exocytosis.
B. Clathrin-coated pit in fibroblast plasma membrane, drawn from a preparation visualised by freeze-fracture, deep etching and rotary replication (original photograph courtesy of Dr John Heusen, in Keen, 1985).
C. Models for the packing of clathrin triskelions in 5- or 6-sided polygon structures of coated pits (Keen, 1985).

may be low; later, they tend to fuse with one another to form structures up to 600nm in diameter in which pH is close to neutral. They move by saltatory motion along tracks of microtubules. Ultimately, after an independent existence of 5-60 minutes, they fuse with elements of the Golgi system. One type of cell that does not have coated pits and appears not to carry out endocytosis is the mature enucleate mammalian erythrocyte. However, all other eukaryotic cells exhibit the endocytotic process continuously except for interruption during mitosis.

The endocytotic pathway of the insulin-receptor complex in particular starts from points of random distribution across the cell surface, proceeding to coated pit locations accomodating a mixture of ligand-receptor types (Maxfield *et al*, 1978). Interestingly, the coated pit structures have been observed clearly only in cultured cells; their involvement *in vivo* has been difficult to establish (Bergeron *et al*, 1985). Subsequent endosomal processing has been investigated by numerous sophisticated biochemical procedures and the large body of information resulting from these studies forms the basis of reviews by Heidenreich and Olefsky (1985) and Sonne (1988), the former dealing with the fate of the receptor, and the latter the fate of insulin. These authors have suggested hypothetical models which unify as much of the available data as possible. The models are described below and summarily represented in Figure 1.7.

Insulin and its receptor dissociate from one another almost immediately in the primary endosome. The drop in pH encountered there does not influence dissociation itself, but it greatly diminishes the reassociation of dissociated ligand, and because the rates of association and dissociation are extremely fast, free insulin quickly accumulates in the vesicle lumen. Shortly after endosome formation, a specific protease, possibly activated by low pH, attacks insulin at one or just a few peptide linkages. The molecular size of the "polypeptide" is

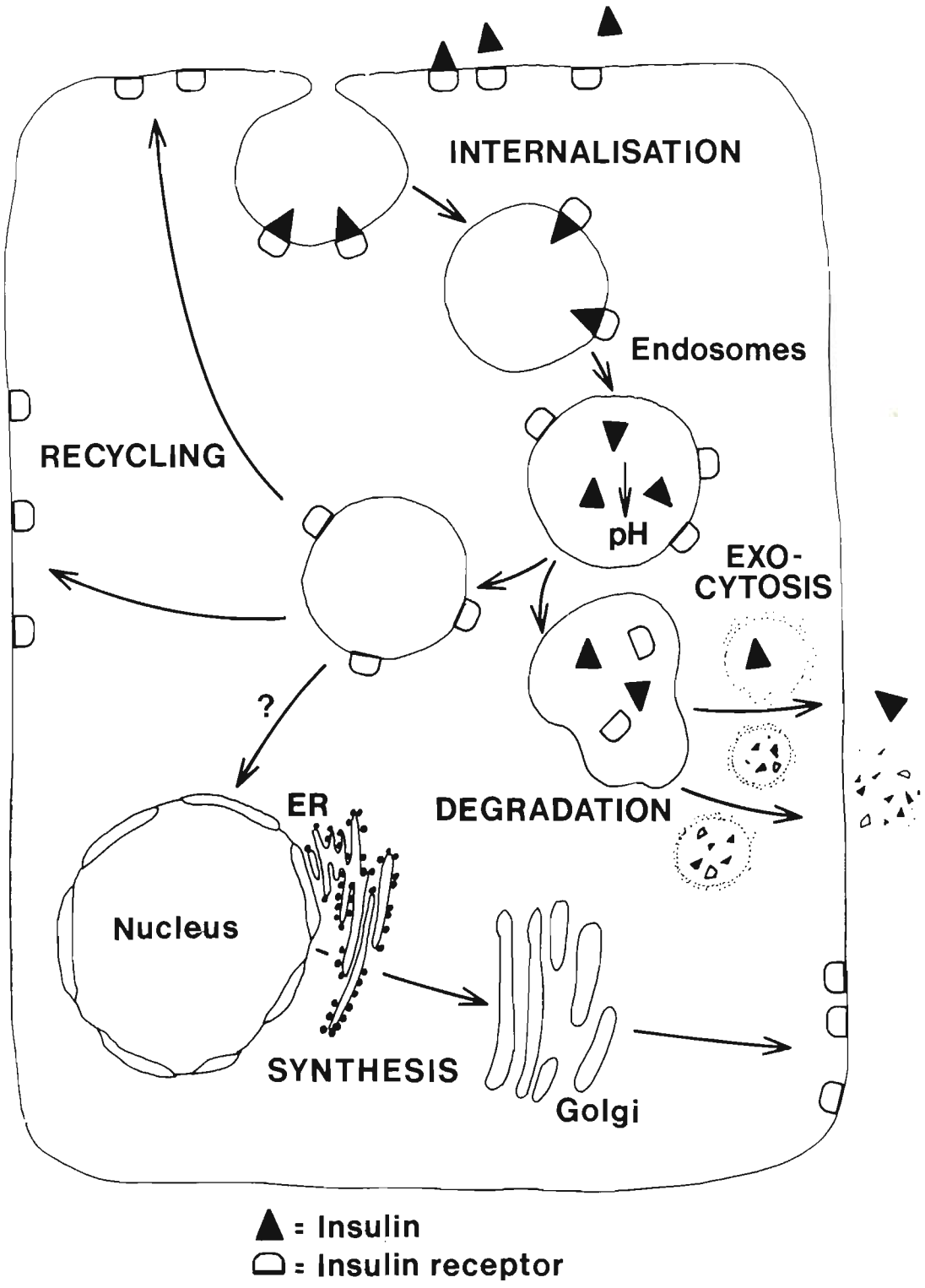


Figure 1.7 Schematic representation of the fate of occupied insulin receptors in target cells (based on Heidenreich and Olefsky (1985) and Sonne (1988).

not affected, but binding affinity is reduced even further. At this stage some intact and "nicked" insulin may be returned to the exterior medium by fusion of endosome with plasma membrane. The remaining "nicked" insulin is further hydrolysed to amino acid status by a protease present in the endosome which is active at neutral pH. Almost all of the internalised insulin is degraded in this fashion. The breakdown products are rapidly released into the extracellular environment; they do not accumulate inside the cell. The physiological explanation of degradation has been the topic of considerable speculation, and is regarded as either (i) that it brings about a rapid quenching of the hormone signal, or (ii) that it is an integral part of the hormone-receptor mechanism.

By contrast with insulin, the internalised receptor has one of three possible fates: it may be recycled back to the cell surface, degraded, or sequestered undegraded in the cell for some time. Most receptors enter the recycling route, but the relative magnitudes of the three route options are varied in order to modulate the final concentration of total cellular receptors, as well as the relative proportions occurring on the surface and in the interior. *Downregulation* is the term applied to the way in which cells exposed to high concentrations of insulin bring about a shift to increased receptor degradation and reduced receptor recycling while the rate of receptor synthesis remains constant; this results in a decrease in the number of surface receptors, and in turn leads to decreased sensitivity to insulin (Gavin *et al*, 1974). Even during downregulation, however, receptor degradation never matches or exceeds insulin degradation. While the population of undegraded intracellular receptors is known to be maintained at a numerically significant level, its subcellular location and specific functional role is not well defined at the present time. Many investigators have suggested a transducing mechanism by which the internalised receptors exert biological effects at specific subcellular sites. Of particular interest in this connection is the reported binding of insulin to receptors in

nuclear membranes (Vigneri *et al*, 1978)), and direct action of insulin on DNA and RNA synthesis (Goldfine *et al*, 1985).

1.3.5 *ENDOCYTOSIS OF MODIFIED LIGANDS*

The endocytosis of chemically modified ligands has been demonstrated in many instances, but clearly depends upon the process of modification leaving the receptor binding site functionally intact. In fact, as mentioned in connection with the insulin receptor, comparative measurement of the binding affinities of a large number of ligand variants chemically altered at specific amino acid residues is one means by which the receptor-binding region may be mapped. Despite this limitation, it has been shown that ligands which have been considerably modified by conjugation to other molecules may nevertheless undergo successful endocytosis, and this facility has been utilised in various ways to contrive transmembrane transfer of the carried molecular species into specific cell types. The targeted endocytotic delivery of drugs has been demonstrated using the principle of covalent attachment of drug to hormone, for example daunomycin or ouabain to melanotropin (Varga, 1985), and drug to receptor antibody, for example ricin A to anti-transferrin receptor (Domingo and Trowbridge, 1985). Copolymers of α-1,4-glucosidase, albumin and insulin have been shown to be transported into intracellular vesicular fractions of cultured muscle cells and to be cleared from the circulation of rats *in vivo*, inferring endocytotic uptake (Poznansky *et al*, 1984). Liposomes incorporating antibody to LDL have been targeted and endocytosed following non-covalent attachment to malondialdehyde-treated LDL (Ivanov *et al*, 1985); in a similar vein, liposomes carrying covalently linked streptavidin have been taken up by cells following non-covalent attachment to biotinylated ß nerve growth factor (Rosenberg *et al*, 1987). In each of these last mentioned studies, however, the liposomes were used only to demonstrate the principle of the method, not to deliver molecules to the target cells. The potential does

exist for the delivery of a broad spectrum of substances, including drugs and DNA, by this means.

In a more extensive investigation, Gitman *et al* (1985) have demonstrated the usefulness of the ligand-receptor system by actually delivering ricin A and SV40 DNA to cells via reconstituted Sendai virus envelopes. These workers covalently linked insulin to the glycoprotein component of the envelope before reconstitution, and removed the Sendai virus receptors from the recipient hepatoma cells by prior neuraminidase treatment. While this type of transfer is not strictly endocytotic, since it involves a degree of direct fusion of the viral envelopes with the plasma membrane, the work is significant in the present context because it illustrates DNA transfer via insulin-receptor interactions. The work described in this thesis does not involve vehicle-mediation, rather being directed towards the delivery to cells of naked DNA in a soluble system. In this respect, the work of Poznansky and his colleagues, mentioned above, is of comparable importance, since they have clearly shown that the use of an albumin carrier linked both to insulin and the molecule to be transported is a practicable one.

CHAPTER TWO

PREPARATION AND CHARACTERISATION OF *N*-ACYLUREA ALBUMIN AND ITS INSULIN CONJUGATE

2.1 INTRODUCTION

The preliminary stages of the work presented in this thesis were concerned with the preparation of a protein macromolecule which might act as both DNA carrier and receptor-specific ligand.

They comprised

- the synthesis of positively charged *N*-acylurea derivatives of bovine serum albumin (BSA) by treatment of the protein with appropriate water-soluble carbodiimides,

- the covalent conjugation of insulin to unmodified albumin and a suitably modified *N*-acylurea albumin respectively by means of glutaraldehyde-induced cross-linkage, and

- the purification and initial characterisation of the various protein products resulting from these procedures.

Carbodiimides are compounds whose general formula is

$$R - N = C = N - R'$$

where R and R′ may be aliphatic or aromatic (Khorana, 1953). There are a number of water-soluble as well as insoluble carbodiimides. Carbodiimides react with carboxyl groups to produce *O*-acylisoureas which in turn react with nucleophiles such as amino groups. Carbodiimides are capable, therefore, of condensing a carboxyl and an amino group to form a peptide bond (Carraway and Koshland, 1972). Additionally, in the presence of

diamines, adjacent carboxyl groups may be cross-linked via the diamine. Such reactions can be achieved under mild conditions and are relatively fast. As a consequence, carbodiimides, particularly those which are water-soluble, have been used extensively in biochemical coupling reactions (Ji, 1983). For example, they have been employed for the synthesis of small peptides (Sheehan and Hess, 1955), the production of cross-linked hormones for use as tools in receptor binding studies (Parsons and Pierce, 1979), the conjugation of toxins to cell-specific carrier proteins in order to obtain targeted killing responses (Faulstich and Fiume, 1985) and, most extensively in recent years, the attachment of weakly haptenic molecules to carrier proteins for the enhancement of anti-hapten antibody reactivities (Goodfriend *et al*, 1964; Bauminger and Wilchek, 1980).

As an alternative to nucleophilic attack generating covalent linkages between molecules, carbodiimide activated carboxyl groups in the form of *O*-acylisoureas may undergo internal rearrangement to produce stable *N*-acylurea carbodiimide adducts (Carraway and Koshland, 1972; Timkovitch, 1977). This transition is illustrated in Figure 2.1. Even under conditions favouring nucleophile interaction (low temperature, an excess of the amino group donor, and a weakly basic pH at which the amino groups are more reactive), *N*-acylurea derivatives are significant by-products of carbodiimide conjugation procedures, while under conditions of moderate temperature and low pH which suppress coupling they are the predominant products (Carraway and Koshland, 1972; Timkovitch, 1977; Davis *et al*, 1984).

Non-conjugative carbodiimide reaction with protein carboxyl groups is potentially useful. Riehm and Scheraga (1966) employed it in a study of the effect of side chain carboxyl group modification on enzyme activity, and Davis *et al* (1984) have suggested detection of complexes formed between receptors and their carbodiimide-modified ligands via antisera prepared against the appropriate carbodiimide adducts. In the work

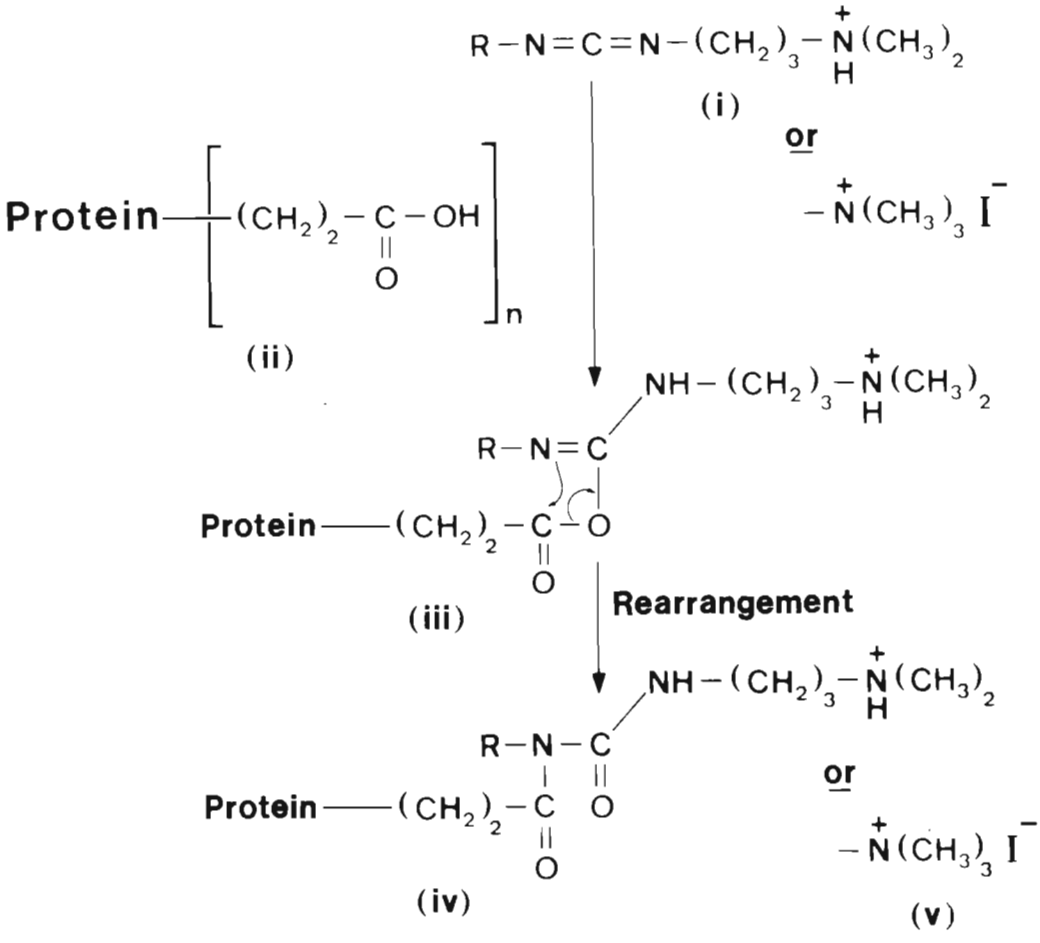


Figure 2.1 Reaction of water-soluble carbodiimide functional groups with protein carboxylic acids under conditions of low pH and moderate temperature

(i) Carbodiimide, either *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide (CDI) or *N*-ethyl-*N'*-(3-trimethylpropylammonium) carbodiimide ($Me^{+}CDI$).
(ii) Protein glutamate residue.
(iii) *O*-Acylurea protein.
(iv) The tertiary *N*-acylurea CDI-protein.
(v) The quaternary *N*-acylurea $Me^{+}CDI$-protein.

described in the following section of this thesis, the carbodiimide induced formation of *N*-acylurea residues at available carboxyl groups on side chains of aspartate and glutamate has been exploited puposefully in order to synthesise protein derivatives which are more electropositive than their native counterparts. This has been achieved by employing water-soluble carbodiimides in which one of the R groups bears a positively charged teriary or quaternary nitrogen function (Figure 2.1) and conducting their reactions with protein (BSA) under weakly acidic conditions (pH 5.5) at 20°C. Two carbodiimides were used: the common *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride, which has been shown to be extremely suitable for carboxyl modification (Timkovitch, 1977), and a derivative of it, *N*-ethyl-*N'*-(3-trimethylpropylammonium) carbodiimide iodide. The latter was synthesised from the former in the laboratory, and was tritiated in the process. Modification of albumin with this carbodiimide thus allowed quantitative determination of *N*-acylurea substitution in addition to introducing into the albumin molecule a quaternary nitrogen function in place of a carboxyl group.

For the coupling of *N*-acylurea albumin to insulin, it was important that the *N*-acylurea protein be rendered free of unreacted carbodiimide prior to conjugation and that a cross-linking agent reactive towards groups other than carboxyls be employed. Accordingly, the modified protein was exhaustively dialysed prior to the coupling, and glutaraldehyde was made the reagent of choice.

Glutaraldehyde has been used extensively for a multiplicity of conjugation procedures. Its mode of action has been the subject of much debate, and a definitive explanation is still lacking despite an important elucidation of the composition of glutaraldehyde solutions and the main features of the chemical properties of these by Monsan *et al* (1975) which has been reviewed subsequently (Peters and Richards, 1977). Several

products may arise during conjugation, and it is still thought likely that a number of unresolved reaction pathways occur (Reichlin, 1980). That this is so is seen by Peters and Richards to be the basis of the success of glutaraldehyde as a non-specific cross-linking agent, and is correlated to the existence of a large number of different polymeric forms of glutaraldehyde in dilute solution. Commercial reagent ("25% glutaraldehyde") has a pH close to 3 and contains a small proportion of glutaraldehyde in equilibrium with its cyclic hemiacetal in both monomeric and polymeric forms (Figure 2.2A). As the pH is increased, however, the glutaraldehyde (a dialdehyde) undergoes an aldol condensation with itself followed by dehydration, generating α,ß unsaturated aldehyde polymers (Figure 2.2B) which continue to polymerise as the pH rises further until precipitation from solution eventually occurs. Despite great variability it appears that most cross-linkages are the result of the interaction of the aldehydic groups of the unsaturated polymers with reactive groups on the conjugant species: ε-amino groups of lysine side chains in the case of protein-protein coupling. Schiff bases are formed. Since in most instances the Schiff base is in conjugation with a double bond (Figure 2.2C), resonance stabilisation occurs and the linkage is resistant to hydrolysis. Proteins reacted with glutaraldehyde thus become partnered in a stable manner through a linker which is itself polymeric and of variable size.

Apart from simplicity of use, efficiency of reaction at room temperature and controllability of cross-linkage, the added advantages of solubility of product and retention of biological activity make glutaraldehyde a favourable choice of coupling agent (Reichlin, 1980). These latter features are especially advantageous in the production of conjugates for use in experiments involving cellular or whole organism processes. It is notable also that in work of a physiological nature involving protein conjugates, serum albumin is the commonest choice of carrier protein (section 1.3.2). Polymerisation of albumin by means of glutaraldehyde has been used in the synthesis of

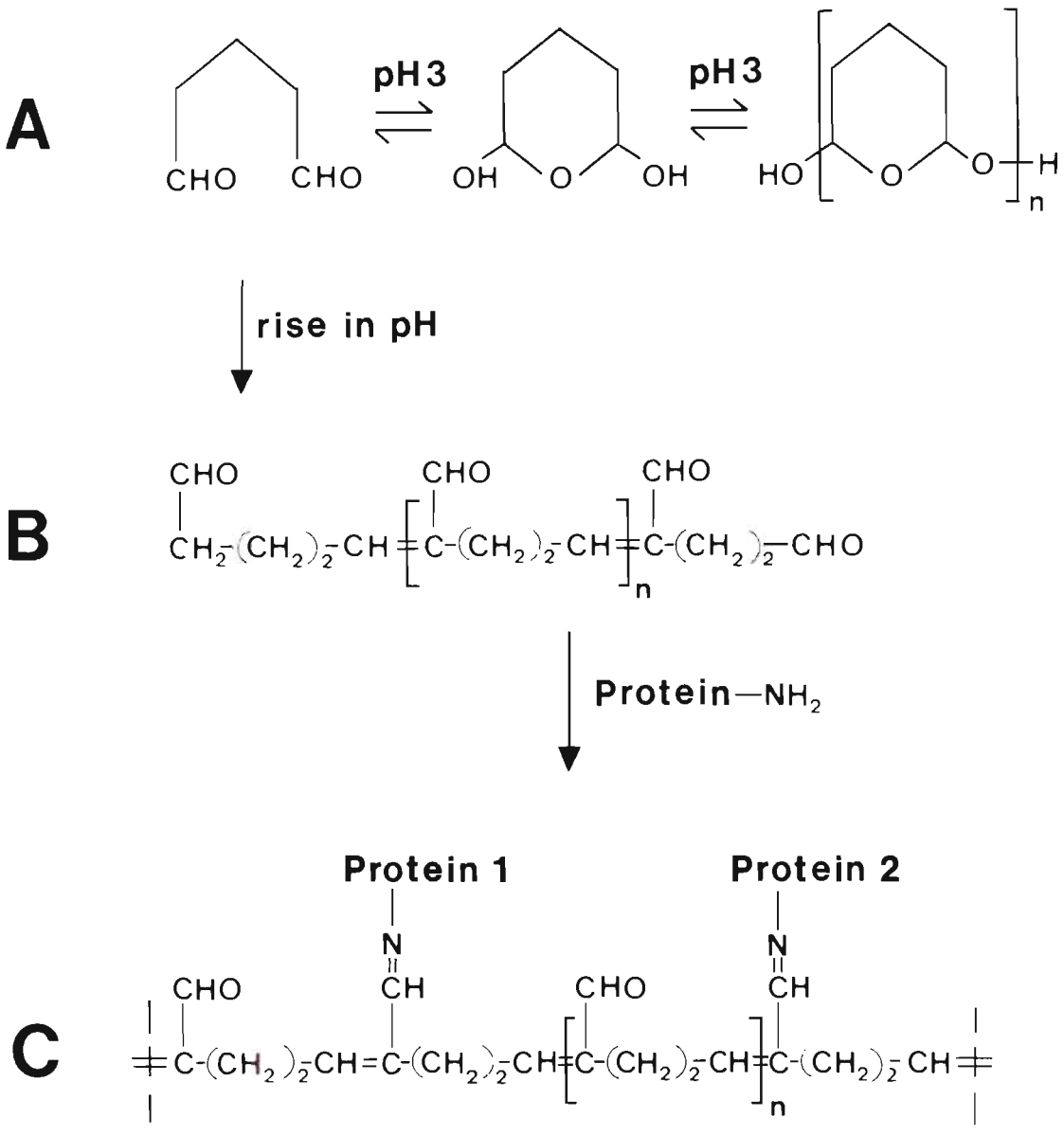


Figure 2.2 The structure and reaction of glutaraldehyde and related compounds (after Peters and Richards, 1977).

A. Forms of glutaraldehyde predominating in"25%" glutaraldehyde (pH 3.1)

B. Forms of glutaraldehyde arising above pH 3 and predominating at neutral pH.

C. Example of glutaraldehyde cross-linkage of 2 proteins by the formation of conjugated Schiff bases between B forms and available ε-NH_2 groups of the protein lysines.

microspheres for drug delivery (Yapel, 1985), the preparation of albumin-L-asparaginase polymers for anti-tumour treatment (Poznansky *et al*, 1982) and the production of albumin-enzyme-antibody complexes for targeted delivery of enzyme to cells (Poznansky and Bhardwaj, 1981). More recent work by Poznansky *et al* (1984) along similar lines describes the glutaraldehyde conjugation of albumin, α-1,4-glucosidase and insulin, and a modification of their method of glutaraldehyde treatment was used in the present work.

Initially reactions were modelled using a relatively crude preparation of bovine insulin rendered zinc-free by column chromatographic purification. Once reaction conditions and the nature of the products had been established, conjugates were made using monocomponent porcine insulin of a quality appropriate to the experimental cell work to follow.

2.2 MATERIALS AND METHODS

2.2.1 *MATERIALS*

Crystalline BSA (Product Code No. A-7638), bovine pancreatic insulin co-crystallised with zinc (approx. 0.5%) (Product Code No. I-5500), ribonuclease A (EC 3.1.27.5) (Product Code No. R-4875) and Cytochrome C (Product Code No. C-7752) were obtained from the Sigma Chemical Co., St. Louis, Mo, U.S.A. Monocomponent porcine insulin (Product No. 7354800) was obtained from Novo BioLabs, Novo Research Institute, Bagsvaerd, Denmark. Anti-porcine insulin (Catalogue No. 5507-1) was purchased from Bio-Yeda, Rehovot, Israel. *N*-Ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide HCl (Code 800907) was purchased from Merck, Darmstadt, Federal Republic of Germany. [^{3}H] Methyl iodide (specific radioactivity 85Ci/mmol) was from Amersham International, U.K. Visking dialysis tubing (20/32) was from Serva Feinbiochemica, Heidelberg, N.Y., U.S.A., and low

molecular weight cut-off dialysis tubing (2000) was from Spectrum Medical Industries, L.A., U.S.A. Acrylamide, *N,N'*-methylenebis-acrylamide, agarose, chemicals for electrophoresis, Bio-Gel A-5m, Bio-Gel P-100 and Immun-Blot kit were obtained from BioRad Laboratories, Richmond, Ca, U.S.A. Sephadex G-50 and G-100 were purchased from Pharmacia Fine Chemicals, Uppsala, Sweden. All other chemicals were of analytical grade.

2.2.2 *METHODS*

Synthesis of [^{3}H] N-Ethyl-N'-(3-Trimethylpropylammonium) Carbodiimide Iodide

The quaternary ammonium carbodiimide *N*-ethyl-*N'*-(3-trimethylpropylammonium) carbodiimide iodide (Me^+CDI) was prepared from *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (CDI) and methyl iodide using the method of Kopczynski and Babior (1984) with minor modifications. In order to radiolabel the product a proportion of the methyl iodide used in the reaction was in the tritiated form. CDI (1.0g, 5.24mmol) was dissolved in 2.0ml water and converted to free base by the dropwise addition of aqueous sodium hydroxide (20%w/v) until the pH of the mixture reached 11-12. The carbodiimide free base released (a colourless oil) was extracted into ether (2 X 20ml plus 1 X 10ml) and the yellow aqueous phase discarded. The pooled ether extract was backwashed with approximately 10ml water and the aqueous layer again discarded. The ether extract was dried over sodium sulphate for 10min, filtered, then incubated at 40°C while ether was removed under a stream of dry nitrogen until only the syrupy free base remained. This was taken up in 15ml dry ether prior to methylation. Methyl iodide (0.62ml, 1.42g, 10mmol) and [^{3}H] methyl iodide (500μCi) mixed with toluene (0.2ml) was added to the ethereal carbodiimide solution with stirring. A creamy yellow coloration and the initiation of crystallisation were apparent immediately. The mixture was kept in the dark at 5°C for several hours to allow

complete crystallisation. Crystals were separated by filtration on a Buchner funnel, washed with dry ether and dried over P_2O_5 *in vacuo*. The hygroscopic product (1.04g, 3.52mmol) had a specific radioactivity of 44μCi/mmol. It was stored desiccated in the dark at -15°C.

Synthesis of N-Acylurea Albumins

BSA was treated with the two water-soluble carbodiimides CDI and Me^+CDI to produce *N*-acylurea CDI-albumin (A-CDI) and *N*-acylurea Me^+CDI-albumin (A-Me^+CDI) respectively. Albumin to carbodiimide mole ratios of reaction were varied widely:

A to CDI - 1:10, 1:50, 1:100, 1:250, 1:500, 1:1000 and 1:2000.
A to Me^+CDI - 1:150 and 1: 500.

In each case a similar reaction procedure was followed. That given below is for a typical treatment of A with CDI at a 1:500 mole ratio.

Aqueous solutions of albumin (14mg, 0.2μmol in 1.3ml) and *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (19.17mg, 100mmol, in 1.5ml) were mixed and the pH immediately adjusted to 5.5 by the addition of dilute HCl. The resultant clear solution was allowed to stand at 20°C for 48-60h in the dark, then dialysed exhaustively against saline-EDTA (50mM NaCl, 0.3mM EDTA, pH 7.0) at 5°C. As in all cases of dialysis the tubing was pretreated (Maniatis *et al*, 1982) and stored under sterile conditions. Aliquots of the dialysed product, *N*-acylurea CDI-albumin (A-CDI) were stored at -15°C.

Protein Determination

Protein was estimated quantitatively by the method of Lowry (1951) using BSA or bovine insulin as standard.

Estimation of the Number of N-Acylurea Groups Attached to Modified Albumins

Aliquots (100µl) of [^{3}H] A-Me$^+$CDI of known protein concentration were precipitated with an equal volume of cold 60%w/v trichloroacetic acid (TCA), filtered through Whatman TF/C filters, washed with 25ml of cold 30%w/v TCA, dried and counted for radioactivity in scintillation fluid (Beckman HP/b).

Determination of Binding Interactions between N-Acylurea CDI-Albumin and Gel Filtration Matrices

Samples of protein (0.6-2.0mg in 1.0ml elution buffer, of known A_{280}) were loaded individually on to columns (50ml burettes) containing Bio-Gel A-5m (agarose), Bio-Gel P-100 (polyacrylamide) or Sephadex G-100 (dextran) and eluted with buffer as indicated in Table 2.2. A series of 3ml fractions representing the entire protein peak were collected, absorption at 280nm measured for each fraction, and recovery of protein calculated as follows:

$$\% \text{ recovery} = \frac{A_{280} \text{ of total recovered protein if in 1ml} \times 100}{A_{280} \text{ of initial protein load}}$$

Where a protein sample was partially retained on a column, the efficiency of removal by an alternative medium was estimated in a similar manner: fractions were collected during elution with an appropriate buffer, absorption at 280nm measured for each fraction, and % recovery calculated.

Preliminary Purification of Bovine Insulin

Prior to use in conjugation reactions, zinc-associated bovine pancreatic insulin was subjected to gel filtration in order to separate the zinc from the peptide. Acetic acid was used as solvent and eluant as advised by Steiner *et al* (1968) in order to ensure, for preparative convenience, that the insulin was

maintained in the monomeric form during filtration. Insulin dissolved in 10%v/v acetic acid (3.0ml per load, 33mg/ml) was filtered through Sephadex G-50 (medium grade) in a 16 X 1100mm column using 10%v/v acetic acid as eluant at a flow rate of 80ml/h. Absorption at 280nm was determined as a continuous trace, and 10ml fractions of peptide material collected manually. Fractions constituting the main insulin peak were pooled and dispensed into glass vials in 3-5ml aliquots. Fractions representing tail-end material were similarly pooled and dispensed but in 10ml aliquots. The insulin solutions were frozen in a thin film round the walls of the vials by rotation in cold ethanol (-60°C) followed by liquid nitrogen (-195.8°C), and were immediately thereafter set up for lyophilisation. The zinc-free insulin product was a fine, fluffy white powder which was stored under desiccating conditions at -15°C. Recovery was 85-90%.

Conjugation of Insulin to Albumins

Insulin was cross-linked to either unmodified albumin or *N*-acylurea CDI-albumin at a 10:1 mole ratio by a development of the method of Poznansky *et al* (1984). The *N*-acylurea CDI-albumin used for this purpose had been prepared at a 1:500 albumin:carbodiimide mole ratio. Insulin (4.2mg, 7.37×10^{-7} moles, dissolved in 600µl dimethyl sulphoxide (DMSO) at 37°C) was mixed with either A or A-CDI (5mg, 7.35×10^{-8} moles, in 1ml saline-EDTA) at 20°C then diluted with 4.0ml 50mM acetate buffer, pH 3.6, and the solution cooled to 5°C. Glutaraldehyde (90µl,"25%" standard reagent) was added and the reaction mixture maintained in the cold overnight. Thereafter, crystalline glycine (75mg) was added, dissolved, and allowed to react with excess glutaraldehyde for 2.5h at 20°C. The reaction mixture was then dialysed exhaustively against 10mM acetate buffer, pH 4.0, containing 5%v/v DMSO at 5°C. Low molecular weight cut-off dialysis tubing was used for this procedure initially. The resultant clear, pale straw-coloured solutions were stored in aliquots at -15°C.

Fractionation of Conjugates

Conjugates were purified and characterised by gel filtration through Sephadex G-100 in a 7 X 1500mm column using 10mM acetate buffer, pH 4.0, containing 5%v/v DMSO as eluant at a flow rate of 6.0ml/h. Absorption at 280nm was determined as a continuous trace. Fractions of 1ml were collected and stored at -15°C for up to 8 weeks. For comparative purposes, unconjugated protein constituents were characterised using the same column and elution conditions.

SDS-Polyacrylamide Gel Electrophoresis of Proteins

Albumin, various *N*-acylurea albumins, insulin and insulin-albumin conjugates were analysed for molecular size characteristics by SDS-PAGE using a BioRad Protean 160mm cell system. Preparation of gels, pretreatment of samples and composition of running buffer were according to Laemmli (1970). Electrophoresis was at 30mA for 4-5h.

Agarose Gel Electrophoresis of Proteins

Albumin, various *N*-acylurea albumins, insulin and insulin-albumin conjugates were analysed for comparative electrophoretic mobility by separation in 1.2%w/v agarose gels at neutral pH. A BioRad Mini Sub DNA cell system was used; gel dimensions 100 X 65 X 3mm. Gels were made and run with buffer containing 36mM Tris-HCl, 30mM sodium phosphate and 10mM EDTA (final pH 7.5). Prior to gel application, protein samples (10-20µg, 3-8µl) were mixed with 1/3 the sample volume of loading buffer (50%w/v sucrose, 4M urea, 0.1%w/v bromophenol blue, 50mM EDTA, pH 7.5). Electrophoresis was at 40V for 3h.

Visualisation of Proteins in Electrophoretic Gels

After SDS-polyacrylamide or agarose gel electrophoresis, proteins were stained with Coomassie Brilliant Blue (CBB) using

one of two techniques according to the solubility properties of the particular proteins present. While most proteins, including albumin, are appropriately precipitated in the methanol-acetic acid-water fixation and staining medium in common use (Weber and Osborn, 1969), insulin is soluble in this and must rather be fixed and maintained in 12.5%w/v TCA for staining to be successful. The insulin staining method given is a version of that of Linde *et al* (1983).

Gels carrying only albumin or albumin derivatives

Fixation - 2h in methanol-acetic acid-water (14:1:5, v/v/v).

Staining - 18h in 0.5%w/v CBB in methanol-acetic acid-water (5:1:5, v/v/v).

Destaining - 2-3 days in several changes of methanol-acetic acid-water (5:1:5, v/v/v).

Storage - in destaining solution.

Gels carrying insulin or insulin conjugates

Fixation - 30min in 10%w/v TCA.

Staining - 18h in 0.025%w/v CBB in 10%w/v TCA.

Destaining - several hours in 10%w/v TCA, 5%v/v acetic acid.

Storage - in 5%v/v glycerol.

Stained gels were photographed by a combination of incident and transmitted white light using Kodak Tri-X Pan black and white film (ASA 400) in a single lens reflex camera fitted with a 55mm close-up lens and a red filter.

Identification of Insulin in the Conjugates

Insulin was detected semi-quantitatively in the conjugate fractions by immuno-dot-blotting. Each fraction tested was diluted in elution buffer to a concentration of 66ng protein/µl. These starting solutions, together with a range of double dilutions derived from them, were applied in 1µl aliquots to

nitrocellulose and processed according to the BioRad Immun-Blot procedure using 0.05%v/v Tween-20, 20mM Tris-HCl, 500mM NaCl, pH 7.5 as incubation and washing medium, gelatin (3%w/v) as blocking agent, anti-insulin (1/333) as first binding agent, protein A-horseradish peroxidase (1/2500) as second binding agent, and a freshly prepared mixture of 4-chloro-1-naphthol dissolved in ice-cold methanol (2ml, 0.3%w/v) and hydrogen peroxide (6µl ice-cold 30% solution in 10ml buffered saline) as colour reagent.

2.3 RESULTS

2.3.1 *N-ACYLUREA ALBUMINS*

Range of Reaction Conditions

Bovine serum albumin (A) was modified by means of the two water-soluble carbodiimides *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (CDI) and *N*-ethyl-*N'*-(tri methylpropylammonium) carbodiimide iodide (Me^+CDI) which respectively incorporate tertiary and quaternary nitrogen functions into the albumin structure by substitution at available side chain carboxyl group sites. In all modification procedures the concentration of albumin in the reaction mixture was 5mg/ml, and treatment was at pH 5.5 and 20°C. However, it was seen to be important to obtain a range of derivatives in which the degree of modification varied widely. Such a series of derivatives would then provide the basis for subsequent DNA binding determinations. Accordingly, reaction mixtures were set up in which the number of moles of carbodiimide per mole of albumin was varied from 10 to 2000.

Number of Carbodiimide Modification Sites per Molecule

Albumin was treated with [^{3}H] Me^+CDI at two different mole

CARBODIIMIDE (mole) / PROTEIN (mole) RATIO	pH of REACTION	MOLES *N*-ACYLUREA ATTACHED PER MOLE ALBUMIN	MOLES (SIDE CHAIN CARBOXYL GROUPS) PER MOLE ALBUMIN
500:1	5.5	27	97
150:1	5.5	13	97

Table 2.1 Reaction of [^{3}H]*N*-ethyl-*N'*-(3-trimethylpropylammonium) carbodiimide iodide with albumin

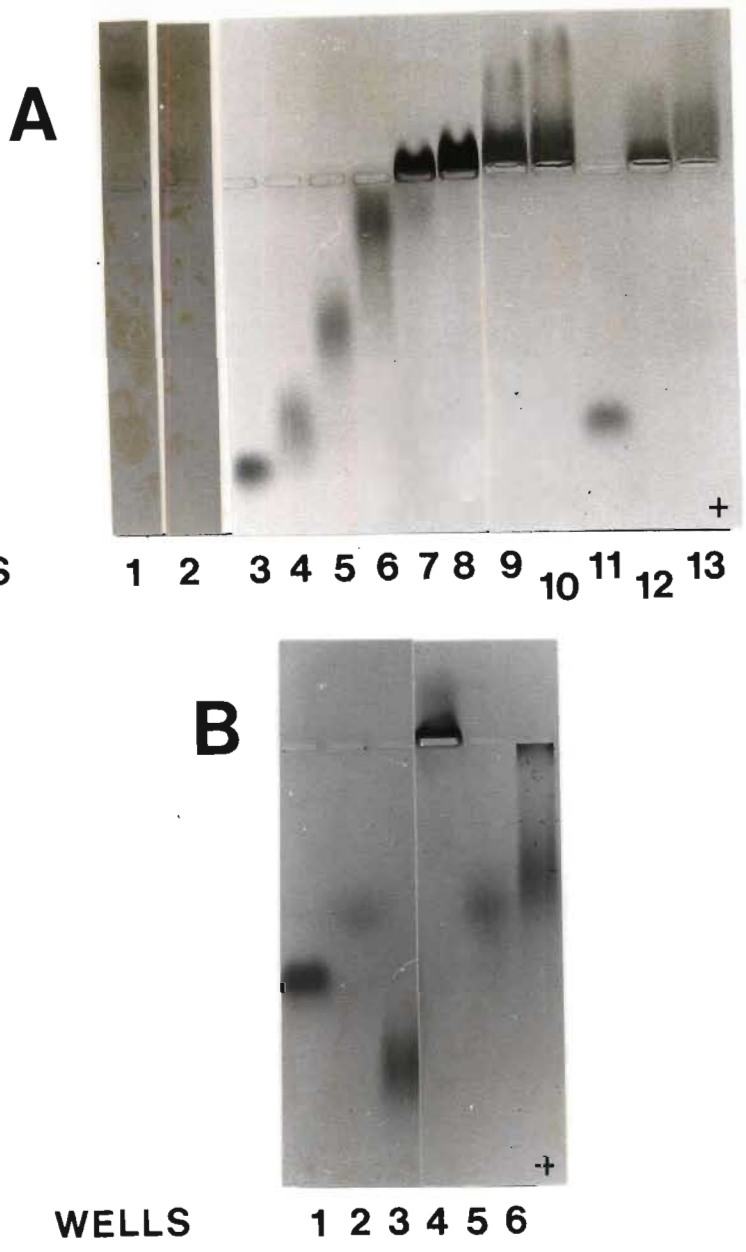


Figure 2.3 Agarose gel electrophoresis at pH 7.5 of proteins and protein derivatives

Modified proteins were dialysed before application to the gel. Protein load, 10-20µg per well.

A. 1, Cytochrome C; 2, RNAse A; 3, Albumin (A); 4-10, *N*-Acylurea CDI-albumins (A-CDI) prepared at protein to carbodiimide mole ratios of 1:10, 1:50, 1:100, 1:250, 1:500, 1:1000 and 1:2000 respectively; 11, Albumin (A); 12 and 13, *N*-Acylurea Me^{+}CDI-albumins (A-Me^{+}CDI) prepared at mole ratios of 1:150 and 1:500 respectively.

B. 1, Albumin (A); 2, Insulin (I); 3, Insulin-albumin (I-A); 4, *N*-Acylurea CDI-albumin (1:500 mole ratio) (A-CDI); 5, Insulin (I); 6, Insulin-[*N*-acylurea CDI-albumin] (I-[A-CDI]).

ratios, 1:150 and 1:500. The use of [^{3}H] Me^{+}CDI served a dual purpose: (i) introduction of *N*-acylurea moieties containing positively charged quaternary nitrogen functional groups into the resultant modified albumins for later use in DNA binding studies, and (ii) determination of the extent of modification (Table 2.1). Albumin contains a total of 97 side chain carboxyl groups on aspartic and glutamic acid residues. Results show that at a mole ratio of reaction of 1:150, 13 of these carboxyl groups became substituted, while at a mole ratio of 1:500 the extent of modification was doubled, with 27 carboxyl groups becoming substituted.

Changes in Net Charge on Modification

Agarose gel electrophoresis of the various carbodiimide modified albumins at pH 7.5 shows that their electrophoretic mobilities differed from that of the parent albumin; they had become more basic (Figure 2.3A). It can be seen that as the A to CDI ratio increased from 1:10 to 1:250 the electrophoretic movement of the A-CDI product towards the positive pole of the gel decreased progressively until it was virtually zero, while at ratios of 1:500 to 1:2000 the A-CDI product moved progressively towards the negative pole of the gel in the opposite direction to unmodified albumin. All these changes indicate increasing electropositivity. Modification by the quaternary carbodiimide Me^{+}CDI effected relatively greater change of net charge than its tertiary counterpart, as demonstrated by a comparison between the mobilities of the respective 1:500 products. The electrophoretic behaviour of the naturally occurring basic proteins cytochrome C and ribonuclease A indicates that their net charge values lie in the same range as those of the various *N*-acylurea proteins synthesised.

Binding to Gel Filtration Matrices

Since *N*-acylurea protein derivatives were considered to have potential as DNA carriers on account of their charge properties,

the corollary existed of their possible electrostatic attachment to gel filtration matrix polymers. Therefore, before gel filtration procedures were seriously initiated for purification of *N*-acylurea protein conjugates, three different polymer matrices were tested for binding interactions. Clearly, the matrix most suitable for subsequent use would be the one demonstrating least binding with *N*-acylurea protein. Table 2.2 summarises the results of tests using gel matrix materials composed of agarose (Bio-Gel A-5m), polyacrylamide (Bio-Gel P-100) and dextran (Sephadex G-100) for elution of A and A-CDI (1:500). Results show that Sephadex, unlike the two Bio-Gel matrices, demonstrated virtually no binding interaction with A-CDI at pH 7.0 (99.5% recovery of protein). Sephadex appeared, therefore, to be the gel filtration medium of choice. When, at a later stage of the work, it became clear that the gel filtration of I-[A-CDI] conjugate necessitated the use of a low pH elution buffer for the maintenance of conjugate solubility at high concentration, investigation into the interaction of Sephadex with *N*-acylurea protein at pH 4.0 was seen to be necessary, since under more acidic conditions the modified protein would bear an even greater net positive charge. Results of these further tests (Table 2.2) indicate that while a small degree of binding was detectable on first passage of A-CDI through the matrix at pH 4.0, 100% recovery of protein was achieved following each repetition of the procedure, suggesting that saturation of binding sites on the matrix had occurred. Ultimately, then, Sephadex G-100 was used for conjugate fractionation (Figure 2.6) under the following conditions: (i) the same eluant (10mM acetate buffer, pH 4.0, containing 5%v/v DMSO) was used for all protein separations; (ii) the column was always 'pre-conditioned' with A-CDI prior to use in the fractionation of A-CDI or its conjugates, and (iii) residual bound material was thoroughly removed by washing with phosphate-buffered 50mM NaCl, pH 7.6, and the column re-equilibrated with low pH eluant prior to use with unmodified proteins or protein conjugates.

Table 2.2 Binding interactions between *N*-acylurea CDI-albumin and various gel filtration matrices

Recovery = % loaded protein eluted from column (50ml)

	ALBUMIN (CONTROL)	***N*-ACYLUREA ALBUMIN (1:500, DIALYSED)**
Biogel A-5m (AGAROSE)	2.0mg sample 100% recovery at pH 7.0	0.6mg sample 47% recovery at pH 7.0
Biogel P-100 (POLYACRYL-AMIIDE)	2.0mg sample 100% recovery at pH 7.0	1.1mg sample 53% recovery at pH 7.0 Attached protein not removed with excess 2M NaCl; fully removed with 1M acetic acid.
Sephadex G-100 (DEXTRAN)	1.0mg sample 100% recovery at pH 7.0	1.1mg samples 99.5% recovery at pH 7.0 (mean of 2 runs: 99% and 100% respectively)
Sephadex G-100 (DEXTRAN)	1.0mg sample 100% recovery at pH 4.0	1.0 - 2.0mg samples 76% recovery at pH 4.0 100% recovery on each repetition on same column at pH 4.0 Attached protein quantitatively removed at pH 7.6

pH 7.0 buffer : phosphate buffered saline (50mM NaCl).
pH 4.0 buffer : 10mM acetate buffer, containing 5%v/v DMSO.
pH 7.6 buffer : phosphate buffered saline (50mM NaCl).

Figure 2.4 is of interest in relation to the results presented above: it shows some binding of A-CDI and A-Me^{+}CDI to an SDS-polyacrylamide gel at the position of the wells, an observation which is in keeping with the findings presented in Table 2.2.

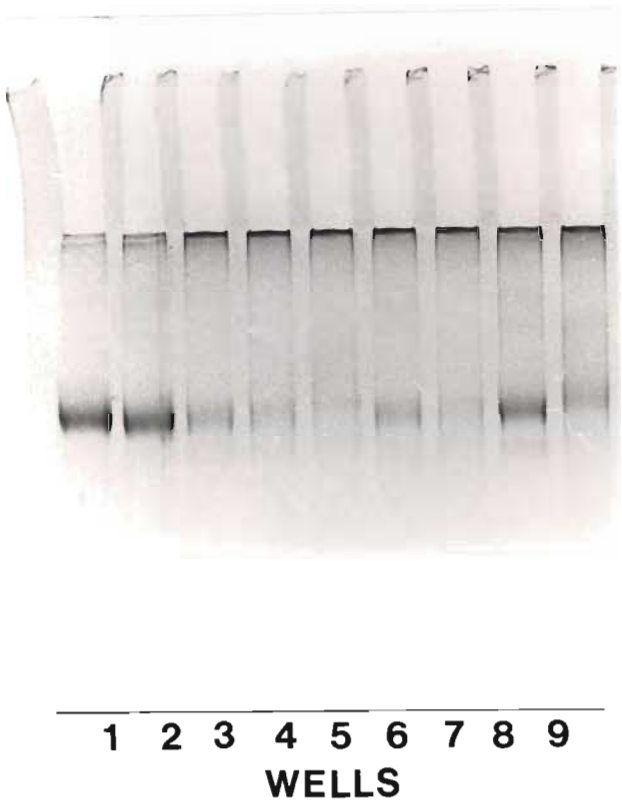


Figure 2.4 SDS-polyacrylamide gel electrophoresis of albumin and its carbodiimide derivatives

12%w/v gel. Modified albumins were dialysed before application, and unless stated were stored at -15°C. Protein load 5µg per well.

1, Albumin (A); 2-5, *N*-Acylurea CDI-albumins (A-CDI) prepared at protein to carbodiimide mole ratios of 1:10, 1:100, 1:500 and 1:1000 respectively; 6 and 7, A-CDI prepared at mole ratios of 1:100 and 1:500 respectively (stored at 20°C for 7 weeks); 8 and 9, *N*-Acylurea Me^{+}CDI-albumins (A-Me^{+}CDI) prepared at mole ratios of 1:150 and 1:500 respectively.

Polymerisation

Size characteristics of *N*-acylurea albumins were investigated by two means: SDS-polyacrylamide gel electrophoresis and gel filtration. Results of SDS-PAGE presented in Figure 2.4 illustrate that (i) *N*-acylurea albumins appeared to be the same size as native albumin except in the case of derivatives produced at the highest carbodiimide concentrations (a protein to carbodiimide mole ratio of 1:2000 in the case of CDI and 1:500 in the case of Me^{+}CDI), (ii) in the products which did appear to be conjugated, the degree of polymerisation shown was slight, and (ii) Me^{+}CDI reaction effected conjugation more readily than CDI, as demonstrated by a comparison between the respective 1:500 mole ratio derivatives. Despite these results, gel filtration through Sephadex G-100 (Figure 2.6A) showed clearly that there was a significant size difference between A-CDI (1:500) and unmodified albumin (A). The latter contained a small proportion of naturally occurring dimer (small peak), and it was at a point equivalent to this that the A-CDI eluted, suggesting that carbodiimide-induced dimerisation had taken place. While gel filtration provides a more sensitive analysis than SDS-PAGE under the conditions used in these experiments, it is possible to combine the findings from both techniques constructively. Overall results then suggest that (i) oligomeric species somewhat larger than dimers might be formed at A:CDI mole ratios of reaction above 1:500 and at even lower A:Me^{+}CDI ratios, but that (ii) the production of large multi-molecular polymers does not occur under any of the reaction conditions used for *N*-acylurea albumin production.

UV Absorption

Figure 2.5 presents ultraviolet absorption spectra of unmodified albumin and A-CDI (1:500) at equal concentrations of protein (500µg/ml; 7.3µM). The two proteins exhibit common peaks of absorption (280nm) and molar extinction coefficients (ε_{280} = 4.4 X 10^{4} $M^{-1}cm^{-1}$). These observations indicate that CDI reaction

did not interfere with the ring structures of the aromatic amino acids responsible for UV absorption. Detection of proteins and their *N*-acylurea derivatives by absorption at 280nm is therefore a straightforward procedure requiring no adjustment for comparison of results to be valid.

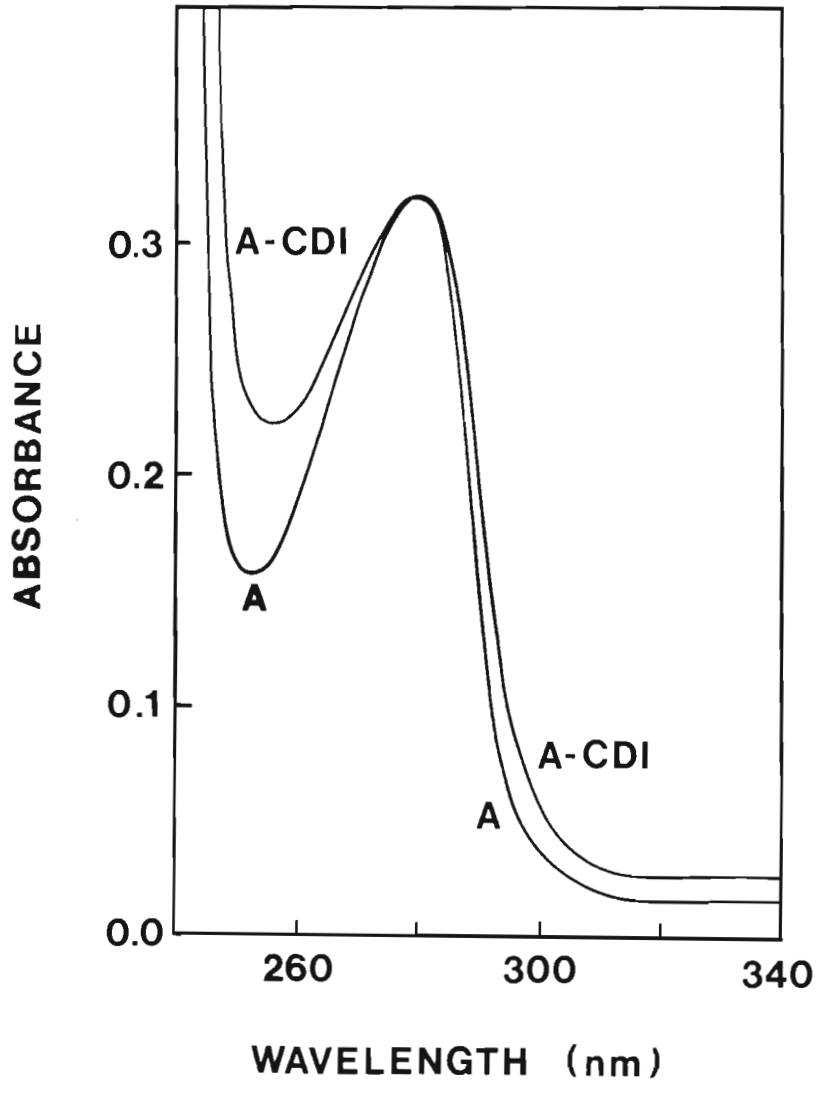


Figure 2.5 UV Absorption spectra of albumin and *N*-acylurea CDI-albumin

Protein solutions, albumin (A) and *N*-acylurea CDI-albumin (A-CDI), each 500µg/ml in 50mM NaCl, 0.3mM EDTA, pH 7.0. The A-CDI was prepared at a 1:500 albumin to carbodiimide mole ratio and exhaustively dialysed after synthesis.

Stability

Results presented in Figure 2.4 suggest that maintenance of *N*-acylurea albumins under ordinary laboratory conditions (20°C) for an extended period of time (7 weeks) did not affect macromolecular integrity by allowing either hydrolytic degradation or molecular aggregation. Since albumin and *N*-acylurea albumin solutions were normally stored at -15°C and defrosted only for sampling, they were assumed to be stable for periods of several months.

2.3.2 *INSULIN-ALBUMIN CONJUGATES*

Composition of Reaction Mixtures

Insulin was covalently cross-linked to unmodified albumin (A) and to *N*-acylurea CDI-albumin (A-CDI) by reaction with glutaraldehyde to provide, respectively, a control conjugate (I-A) and a conjugate carrying *N*-acylurea carbodiimide modifications (I-[A-CDI]). The A-CDI used in the conjugation reaction had been prepared at a 1:500 mole ratio, and therefore possessed 27 N-acylurea groupings (Table 2.1) and considerable electropositivity (Figure 2.3). In order to ensure the attachment of several molecules of insulin (mol wt 5700) to each molecule of albumin (mol wt 68000) reaction mixtures contained an insulin to albumin mole ratio of 10:1 (mole calculations based on unmodified albumin input in both cases).

Molecular Size of Conjugates

As with *N*-acylurea albumins prior to conjugation, size characteristics of conjugated products were determined by means of both SDS-polyacrylamide gel electrophoresis and gel filtration. Results of SDS-PAGE of I, A and I-A are presented in Figure 2.7. They show that the I-A product was a heterogeneous mixture of species ranging in size from only

slightly larger than albumin itself to a molecular conglomerate which moved very slowly, hardly entering the gel. This high molecular weight complex represents a significant proportion of the total conjugate population and forms a heavily stained band within the overall 'smudge'. Respective gel filtration of I-A and I-[A-CDI] through Sephadex G-100 (Figure 2.6) gave rise to identical profiles representing a series of conjugates in which the very high molecular weight components dominate, forming a leading peak which is followed by a shoulder of smaller conjugates. Information concerning molecular size provided by gel filtration is entirely compatible with that provided by SDS-PAGE analysis. The two are correlated directly in Figure 2.7 which shows the electrophoretic mobilities of two gel filtration fractions. Fraction 18, representing the main peak (Figure 2.6), is seen to be equivalent to the discrete high molecular weight band on the gel. Fraction 24, representing the shoulder, appears to consist of products of a size range intermediate between the largest conjugates and unconjugated albumin.

Incorporation of Insulin

Indirect evidence supporting the covalent inclusion of insulin within the high molecular weight conjugates resulting from glutaraldehyde cross-linkage is provided from two sources: (i) SDS-PAGE of I-A reaction mixture containing 46%w/w insulin, dialysed in low molecular weight cut-off tubing before electrophoresis, indicated a disproportionately faint band at the position of unconjugated insulin (Figure 2.7), and (ii) Sephadex G-100 gel filtration of similarly dialysed conjugation reaction mixtures showed only an insignificant amount of protein at the position of unconjugated insulin (Figure 2.6). Direct evidence is provided from a third source: immuno-dot-blot analysis of gel filtration fractions 16 to 25 (Figure 2.8). Results of this analysis demonstrated unambiguously that insulin was present in fractions 17 to 25, the highest ratios of insulin to total protein being apparent in fractions 18 to 24.

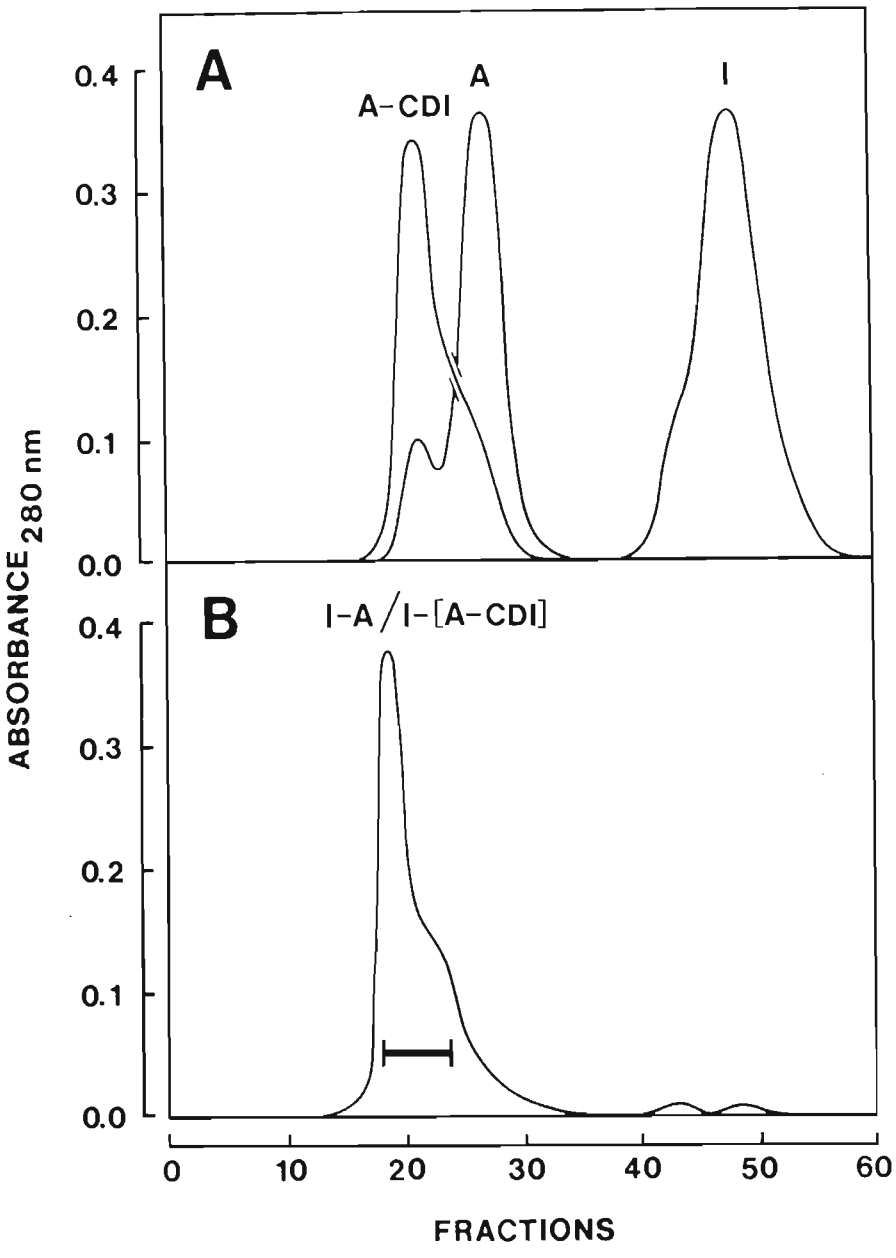


Figure 2.6 Gel filtration chromatography of insulin-albumin conjugates and their constituent proteins

Proteins were purified and characterised by gel filtration through Sephadex G-100 as described in Materials and Methods. Reaction mixtures were dialysed exhaustively prior to loading on to the column; reaction mixtures containing insulin were dialysed in low molecular weight cut-off (2000) dialysis tubing.

A. Unmodified insulin (I), unmodified albumin (A) and *N*-acylurea albumin (A-CDI).

B. Insulin-albumin (I-A) and insulin-[*N*-acylurea albumin] (I-[A-CDI]).

The horizontal bar shows those conjugate fractions containing the highest proportions of insulin to total protein as estimated by immuno-dot-blot analysis (Figure 2.8).

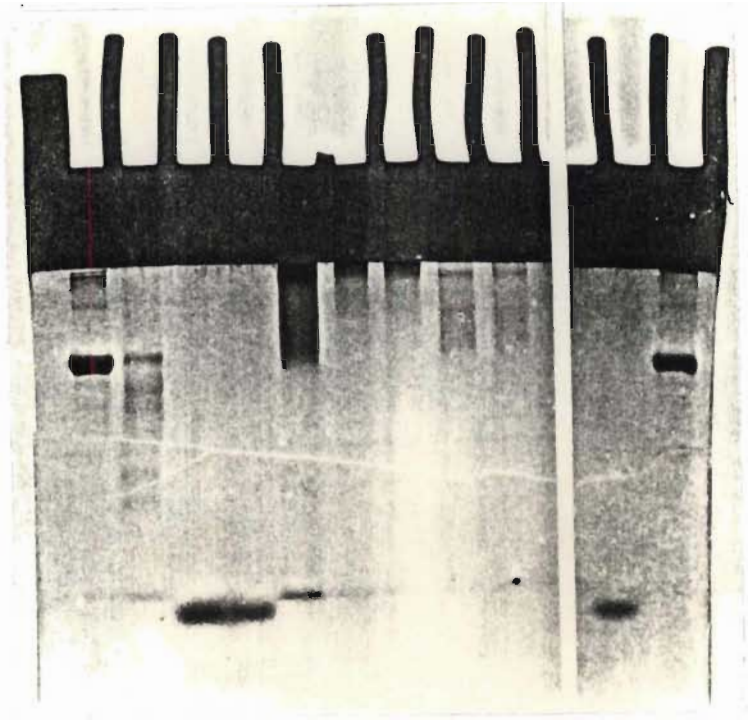


Figure 2.7 SDS-polyacrylamide gel electrophoresis of insulin-albumin and its constituent proteins

4%w/v stacking gel; 12%w/v resolving gel. Protein load 10µg per well.

1 and 2, Albumin (A), untreated and alkali digested respectively; 3, Zinc-associated insulin (Zn-I); 4, Zinc-free insulin (I); 5-9, Insulin-albumin conjugates (I-A): 5, dialysed I-A reaction mixture prior to gel filtration; 6 and 7, Sephadex G-100 I-A fraction 18, untreated and alkali-digested respectively; 8 and 9, Sephadex G-100 I-A fraction 24, untreated and alkali-digested respectively; 10, Zinc-free insulin (I); 11, Albumin (A).

Alkali digestion: sample made 0.2M with respect to NaOH, incubated at 20°C for 2h, then neutralised with HCl before electrophoresis.

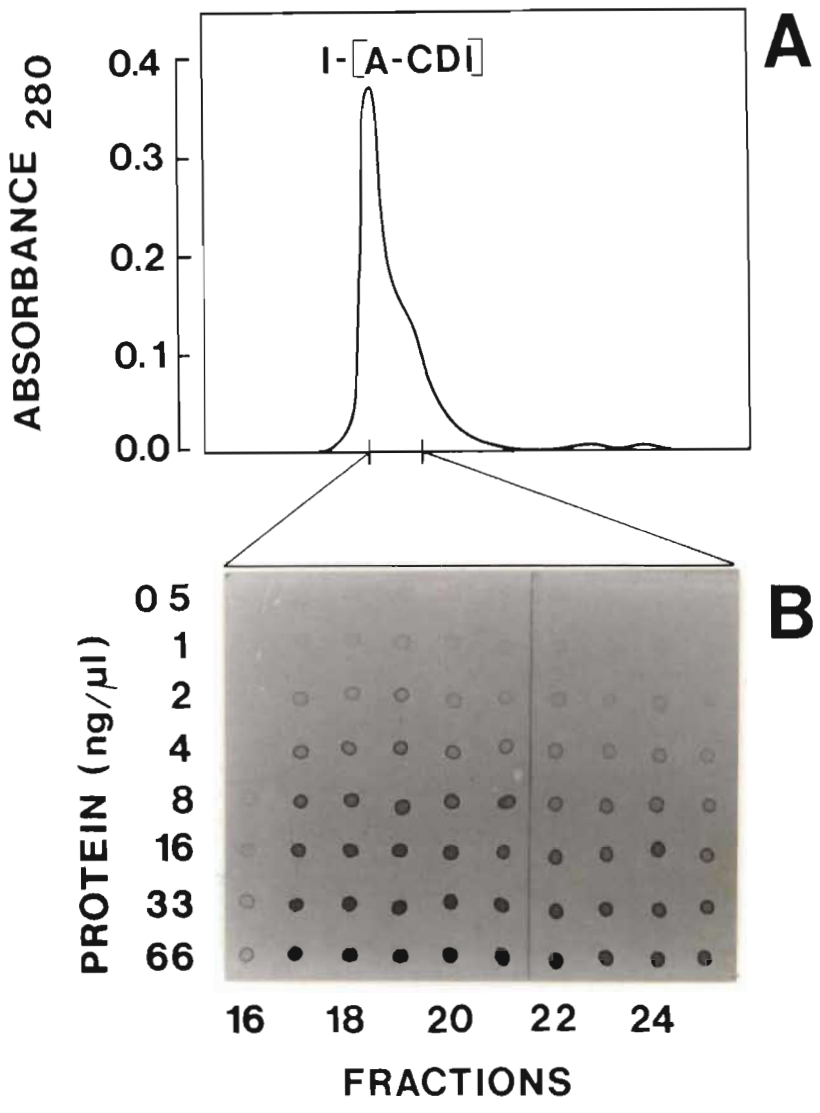


Figure 2.8 Immuno-dot-blotting for the identification of insulin in gel filtration fractions of insulin-[*N*-acylurea CDI-albumin]

A. Sephadex G-100 gel filtration profile of insulin-[*N*-acylurea CDI-albumin] (I-[A-CDI]).

B. Dot blots (1µl) of fractions 16-25, each normalised to a starting concentration of 66ng/µl and thereafter serially diluted to 0.5ng/µl. Insulin was detected in the blots by successive incubations with (i) anti-insulin, (ii) Protein A-horseradish peroxidase, and (iii) a mixture of 4-chloro-1-naphthol and hydrogen peroxide, which gives rise to a purple coloration in the presence of the peroxidase enzyme.

Net Charge

The charge characteristics of insulin-albumin conjugates and their component proteins are illustrated in Figure 2.3B. Notably, I-A appears to be more acidic than either of its constituent proteins, presumably because many of the basic lysine ε-amino groups contributing to overall charge of the native proteins have been utilised in glutaraldehyde cross-linkages. A comparable shift towards electronegativity is shown in the I-[A-CDI] conjugate, but since the A-CDI constituent prior to conjugation is much more positively charged than native albumin on account of its modified aspartate and glutamate residues, the resultant conjugate is significantly more electropositive than I-A.

Stability

Results presented in Figure 2.7 indicate that the molecular size of conjugates remained unaffected by conditions of mild alkali digestion which did cause some degradation of unconjugated albumin. This observation confirms the stable nature of the covalent glutaraldehyde-based intermolecular linkages holding the conjugant species together. In addition, it suggests that the tertiary structure of the component proteins might have become stabilised by covalent intramolecular linkages.

2.4 SUMMARY AND DISCUSSION

Bovine serum albumin has been reacted with water-soluble carbodiimides at various mole ratios under conditions known to disfavour conjugative protein-protein coupling but to bring about substitution of carboxyl groups with *N*-acylurea carbodiimide adducts. It has been shown that *N*-acylurea albumins resulting from this type of modification

- may undergo dimerisation or even greater degrees of oligomer formation, but do not become polymerised into large molecular weight complexes (Figure 2.4);

- retain radioactive label on reaction with *N*-ethyl-*N'*-(trimethylpropylammonium) carbodiimide iodide which has been tritiated at one of the methyl groups, indicating incorporation of the the trimethylpropylammonium moiety into the protein structure (Table 2.1);

- exhibit charge properties which are altered towards greater electropositivity in a manner explicable by substitution of side chain carboxyl groups with tertiary and quaternary amino group functions characteristic of the reacting carbodiimide (Figure 2.3A), and

- display a degree of carboxyl group modification, as measured by either incorporation of radioactivity or change in electrical charge, which is directly related to the concentration of carbodiimide in the reaction mixture (Table 2.1; Figure 2.3A).

Reaction of albumin with *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (CDI) was undertaken at seven different mole ratios varying from 1:10 to 1:2000, and it has been shown that these conditions effect measurable and progressive changes of charge in the products (Figure 2.3A). At mole ratios up to 1:250 the degree of charge change brought about by each stepped increase in carbodiimide concentration is very marked. At mole ratios of 1:500 and above, the changes resulting from even larger increases in carbodiimide input are relatively minor, suggesting that under these conditions readily accessible reaction sites are reaching limiting numbers. It has been established from modification with [^{3}H]Me^{+}CDI that 13 carboxyl groups react at a mole ratio of 1:150 , while 27 groups react at a ratio of 1:500. It can be tentatively suggested, then, that the number of reaction sites existing on the exposed

periphery of the albumin molecule under physiological (non-denaturing) conditions is in the region of 30 or so.

Part of the rationale for the use of the two different carbodiimides in synthesising *N*-acylurea proteins was that it would be interesting and possibly useful to compare the effects of tertiary and quaternary nitrogen groupings on DNA binding capacity: it was thought that the quaternary compounds might bind DNA more effectively than the tertiary equivalents, since they would exert a stronger charge. Although albumin was modified with *N*-ethyl-*N'*-(trimethylpropylammonium) carbodiimide iodide (Me^+CDI) at only two different mole ratios, one of them (1:500) was selected to provide direct comparison of the product with that of A:CDI reaction at the same mole ratio. Results of agarose gel electrophoresis have indicated that A-Me^+CDI is indeed more basic than the equivalent CDI product (Figure 2.3A).

Of the various *N*-acylurea albumins prepared and analysed in subsequent DNA binding experiments (Chapter 3), the reaction products prepared at a 1:500 mole ratio were the ones eventually selected for glutaraldehyde cross-linkage to insulin. Choice was based upon the fact that these compounds, already shown to exhibit a significant degree of modification (Table 2.1) and concomitant electropositivity (Figure 2.3), proved to have a marked DNA binding capacity (Figures 3.1 and 3.4). Ultimately, however, the unfavourable solubility characteristics noted in derivatives of the A-Me^+CDI product led to the sole use of A-CDI as conjugant.

Glutaraldehyde reaction with mixtures of insulin and unmodified albumin or insulin and A-CDI has been shown here to give rise to products which

- comprise a heterogeneous population of conjugates ranging widely in size, in which the majority are several times larger than unconjugated albumin (Figures 2.6 and 2.7);

- show the covalent incorporation of insulin into their structures (Figure 2.8), and

- demonstrate some loss of electropositivity compatible with cross-linkage via lysine ε-amino groups (Figure 2.3B).

The glutaraldehyde reaction mixture contained a large excess of insulin molecules over albumin molecules (mole ratio of 10:1; 46%w/w insulin) in order to create the probability of cross-linking the proteins in proportions of the same order. In synthesising conjugates consisting of three different proteins including albumin and insulin, Poznansky *et al* (1984) have used a similar strategy, and appear confident in interpreting conjugate composition on the basis of relative molar contribution. While such an assumption may not be precisely valid, it is clear from several sources of evidence in the work presented here (Figures 2.6 - 2.8) that virtually all of the insulin in the original mixture became incorporated into conjugated species larger than monomeric albumin, thus suggesting that heterologous coupling had occurred in approximately the proportions planned. On SDS-PAGE (Figure 2.7), most of the conjugates appeared to occupy a band position extremely close to the well of origin and some distance from the albumin control: sufficiently immobile to represent a molecular weight of several hundred thousand. A complex consisting of three albumin molecules combined with thirty insulin molecules (molecular weight, 375000) is therefore one speculative example out of many possible products. Determinations of precise molecular weights and composition were not undertaken; the extreme heterogeneity of the conjugate population was a major reason for this.

Having modified albumin with carbodiimide in order to produce basic proteins which might be capable of electrostatic binding to DNA, and having established that a binding affinity for DNA does exist in these compounds (Figures 3.1 and 3.2), the partial reverse shift in charge accompanying their glutaraldehyde

conjugation to insulin (Figure 2.3) seemed to be a retrogressive step. However, it was clear that there existed a significant net difference in charge between the I-A control conjugate and the experimental I-[A-CDI] (Figure 2.3B), the latter being considerably more electropositive. The possibility existed, therefore, that some capacity for DNA binding would be shown by this conjugate. Its DNA binding potential is reported in the following chapter in conjunction with the preliminary and subsequent studies of DNA binding by unconjugated albumins.

Bovine and porcine insulins possess only one lysine residue, at position B29 (Smith, 1972). If glutaraldehyde action does in fact occur only at lysine side chain amino groups, insulin must become bridged to albumin at this one point. The fact that the lysine is the penultimate residue in the B chain and is situated on the periphery of the molecule (Figure 1.4) suggests the probability that normal structure is retained in the major portion of the peptide following glutaraldehyde reaction. In such a circumstance retention of conformational integrity and specific binding affinities would be likely. Residue B29 itself is not involved in receptor binding (Gammeltoft, 1984) (Figure 1.4). These theoretical considerations support the concept of a glutaraldehyde-linked insulin-[*N*-acylurea albumin] conjugate acting as an effective cell-specific ligand as well as a DNA carrier. Experimental work designed to investigate conjugate-cell receptor binding is presented in Chapter 4.

CHAPTER THREE

BINDING OF DNA TO *N*-ACYLUREA ALBUMIN AND ITS INSULIN CONJUGATE

3.1 INTRODUCTION

Once the various *N*-acylurea derivatives of albumin had been obtained and purified (Chapter 2), they were subjected to DNA binding assays in order for their potential as DNA carriers to be determined. The preliminary results of these assays led to the choice of *N*-acylurea albumins prepared at albumin to carbodiimide mole ratios of 1:500 being selected for conjugation to insulin (Chapter 2). Of the alternatives, *N*-acylurea CDI-albumin (A-CDI) and *N*-acylurea Me^{+}CDI-albumin (A-Me^{+}CDI), the former proved to demonstrate more suitable solubility properties (section 2.4), and the resultant conjugate, I-[A-CDI], was in turn assayed for DNA binding capacity using the same methods. In addition, experiments were undertaken to investigate the nature of the binding interactions involved.

The central importance of protein-DNA binding relationships in genetic structures and processes has been brought into sharp focus over the past two decades. The formation of protein-DNA complexes is now seen to be fundamental to the architectural and functional integrity of genetic material; it is only by means of the activity of proteins that genetic messages are conserved, identified and acted upon during growth, development and reproduction. Binding proteins play roles in chromosome conformation (Pederson *et al*, 1986), nuclear scaffolding (Newport and Forbes, 1987; Razin, 1987), DNA replication (Jackson and Cook, 1986), genetic recombination (Cox and Lehman, 1987; West, 1988), DNA repair (Sancar and Sancar, 1988), regulation of gene expression via transcriptional initiation

(Ptashne, 1986; Melton, 1987) and nuclear RNA processing and transport (Mariman *et al*, 1982). The relevance of protein-DNA association to such a wide range of core activities in the cell has meant that several appropriate binding assays have been developed.

In one of the first investigations leading to successful detection and purification of a bacterial repressor, Ptashne (1967) used a zonal sedimentation technique, and this approach is still employed effectively for the identification of binding proteins in experiments which do not demand high sensitivity or resolution and for which the protein is relatively plentiful. For example, Mignotte *et al* (1985) have utilised glycerol gradients for the characterisation of a single-stranded DNA-binding protein having a possible role in mitochondrial nucleoid replication, and Kuhn *et al* (1987) have used sucrose gradient analysis of intermediate filament protein interaction with supercoiled DNA to advance an argument concerning the potential role of vimentin and related filamentous proteins in the control of DNA replication and expression.

Several techniques have been devised for the recovery of protein-DNA complexes from crude extracts, and such methods may sometimes be used for assay purposes. One is the immunoprecipitation of complex with antibody raised against the binding protein (Lee and Knipe, 1986). Another with comparable applications is DNA affinity chromatography, a method which is used mostly for preparative work but which has found occasional analytical use (Gaudray *et al*, 1981). A more recently developed technique is that of protein-DNA blotting, in which a mixture of proteins separated by polyacrylamide gel electrophoresis is bound on to nitrocellulose by blotting, incubated with a radioactively labelled DNA probe, washed with varying degrees of stringency, and autoradiographed for visualisation of the binding proteins. This approach has yet to be refined; its uses to date have involved the disadvantage of protein denaturation by SDS (Bowen *et al*, 1980) or low pH (Levinger and Varshavsky,

1982). Modifications employing non-denaturing electrophoretic systems have not been reported.

Two techniques much used for the finer analysis of protein-DNA binding reactions are the filter binding assay and the electrophoretic mobility shift assay respectively, both of which are suitable for use with semi-purified or purified protein components, and are extremely sensitive. These methods were employed in the present work to demonstrate and measure the extent of binding of *N*-acylurea albumin compounds to two forms of DNA: linear DNA fragments of heterogeneous sequence from calf thymus, and the circular DNA of purified pBR322 plasmid. In addition, the filter binding assay was used to investigate various parameters of the binding reaction in order to help clarify the nature of the forces operating in the protein-DNA complex.

The principle of the filter binding assay rests on the fact that proteins, and therefore also protein-DNA complexes, are retained by nitrocellulose membranes upon filtration, whereas double-stranded DNA fails to be retained in this manner under the same conditions. Thus if a binding protein is preincubated with radioactively labelled DNA, filtered through nitrocellulose, and the filter washed with buffer and processed for counting, the amount of DNA bound may be quantified. The assay has been used a great deal for a variety of studies. It has been the means of measuring the kinetic and equilibrium parameters of the binding interactions of purified proteins with DNA (Riggs *et al*, 1970), as well as the rapid monitoring of binding during purification of a binding protein (Johnson *et al*, 1980). In a secondary role, it has been used for the detection of nuclear binding proteins in crude extracts, and, in conjunction with restriction endonuclease cleavage, identification of specific fragments of DNA carrying the binding sites has been possible (Borgmeyer *et al*, 1984). However, the method has serious limitations as a general detection system as it can demonstrate the activity only of high affinity DNA

binding proteins (Hennighausen and Lubon, 1987). In the present work involving examinination of the binding of synthesised and purified *N*-acylurea proteins to DNA, the filter binding assay was found to be particularly precise and versatile, lending itself to both quantitative and qualitative use.

The mobility shift assay, for which the term "band shift assay" has been suggested (Varshavsky, 1987), is a more recent alternative to the filter binding assay. It is based on differences in size leading to separation of protein-DNA complexes from free DNA and from each other during gel electrophoresis. A version of the technique was originally introduced for the separation of nucleosomes formed *in vivo* (Varshavsky *et al*, 1976), its use for the assay of complexes formed *in vitro* being initiated later by Fried and Crothers (1981). Most band-shift assays have employed polyacrylamide gel electrophoresis, but agarose gel electrophoresis has been introduced recently (Berman *et al*, 1986; Huckett *et al*, 1986) and was used in the work presented in this chapter. As with the filter binding assay, the band shift approach to kinetic studies preceded its application to crude nuclear extracts, but its usefulness in this latter area has been accented of late, mainly because of the extreme sensitivity of the method. Hennighausen and Lubon (1987) state that its sensitivity exceeds that of the filter binding assay by at least an order of magnitude; Varshavsky (1987) has suggested that the exceptional potential for sensitivity may be due partly to stabilisation of the protein-DNA complexes once they enter a gel; the gel matrix may exert a "caging" effect thus decreasing the rate of dissociation. However, extreme sensitivity probably exists only under optimal conditions, which, for naturally occurring binding proteins, include the use of small DNA fragments (less than 300bp) and buffers of low ionic strength (10mM) (Hennighausen and Lubon, 1987; Varshavsky, 1987). In the work reported here, the DNA used was of moderately high molecular weight, in keeping with the large sizes of the synthesised *N*-acylurea binding proteins, and gels were run under normal buffering conditions.

The band shift assay was regarded, therefore, as a means of generating supportive evidence for binding, not as a primary, finely tuned analytical tool.

In vitro binding assays have generated definitions of some of the important conditions required for protein-DNA complex formation to take place, and in doing so have pointed to possible modes of interaction at the binding site. A consistently observed characteristic of protein-DNA binding is a marked salt dependency (Record *et al*, 1976; Karpel *et al*, 1981; Stahl *et al*, 1984) indicative of electrostatic interaction between basic amino acids of the protein and phosphate groups of the nucleic acid. Since *N*-acylurea albumins are modified proteins bearing positively charged tertiary amino or quaternary ammonium groups attached via a spacer to aspartic and glutamic acid side chains, it seemed likely that the formation of salt bridges between these groupings and ionised phosphodiester backbone regions of DNA would constitute one component of the binding response. Therefore the effect of sodium ion concentration on *N*-acylurea protein-DNA binding was investigated. Furthermore, since electrostatic binding is implicitly charge dependent, the effect of pH on binding was also examined.

Out of protein-DNA binding assays in general have developed competition binding assays of various kinds. Many of these are for use in identifying binding proteins from crude extracts, and are founded on the idea of introducing into the binding mixture heterologous DNA fragments (Riggs *et al*, 1970; Borgmeyer *et al*, 1984) or synthetic alternating copolymer duplexes (Singh *et al*, 1986) which decrease the relative contribution of nonspecific protein binding to a specific DNA probe. A less common kind of competition assay, but one of relevance in this work, involves the use of the polyanion heparin to compete with DNA for binding to a specific binding protein as a means of analysing the relative contributions and strengths of ionic and nonionic components of binding conducted under different conditions

(Melancon *et al*, 1982). The same strategy was used here to study the ion dependency and tenacity of the *N*-acylurea protein-DNA bond in relation to time elapsed after formation.

Some biochemical techniques directed towards detailed sub-molecular analysis of protein-DNA binding interactions, not employed in the present investigation, are nevertheless of considerable relevance because of the conceptual advances made as a consequence of their use. "Footprinting", by which precise binding site sequences in the DNA may be mapped, applicable both *in vitro* (Hennighausen and Lubon, 1987) and *in vivo* (Jackson and Felsenfeld, 1987), has proved to be one of the more important of these methodologies for the study of the specific contribution of DNA to the binding reaction. It is now known that DNA is recognised by binding proteins in a non-specific or specific manner. Proteins binding non-specifically show poor base sequence recognition but may bind preferentially to double-stranded or single-stranded forms of DNA. This category includes the DNA melting proteins such as gene 32 protein of bacteriophage T4 (Jensen *et al*, 1976) and bovine pancreatic ribonuclease (Jensen and von Hippel, 1976); also some proteins, like recA, which bind both single-stranded and duplex DNA but in different modes (West, 1988). Many regulatory proteins which act at specific genomic target sites also exhibit general affinity for other regions of DNA. Thus the binding to DNA of RNA polymerase holoenzyme from *E. coli* (Melancon *et al*, 1982) and catabolite gene activator protein (CAP) from *E. coli* (Weber and Steitz, 1984) are interactions which include non-specific responses, even though the more characteristic binding activities are highly site-specific.

Detailed investigations into the protein components of binding have dwelt upon those proteins involved in site-specific interactions. They have included genetic studies of mutant proteins defective in DNA binding (Hecht *et al*, 1983), and X-ray crystallographic analyses of DNA binding proteins such as Lambda Cro (Anderson *et al*, 1981), CAP from *E. coli* (McKay and Steitz,

1981) and Lambda repressor (Pabo and Lewis, 1982). Such approaches eventually led to the proposition of a common model for the physico-chemical interactions occurring at the binding sites of naturally evolved site-specific protein-DNA complexes (Pabo and Sauer, 1984). Evidence for the validity of this model has come more recently from X-ray crystallography of a co-crystal of the DNA-binding region of the repressor protein of coliphage 434 and the short (14bp) sequence that the protein recognises (Anderson *et al*, 1985), as well as studies of the interactions of protein-engineered 434 repressor variants (Wharton and Ptashne, 1985). It has also found general support from sources such as electron microscopic visualisation of protein-DNA complexes (Stasiak *et al*, 1981; Kuhn *et al*, 1987) which has been an alternative means of 3-dimensional image reconstruction (Egelman and Stasiak, 1986).

The model of interaction between DNA binding proteins and DNA is briefly as follows: The binding protein is usually a dimer or tetramer with symmetrically related subunits which carry a characteristic motif of two α-helices linked by a tight turn. One of the α-helices protrudes from the surface of the molecule. The complete protein thus has two or more projecting α-helical regions. The angles at which these helices are borne, their distance apart and their diameters (including side chains) are such that they fit neatly into successive major grooves along one face of double stranded DNA in B conformation, thereby acting as backbone clamps from which amino acid side chains can contact the edges of the base pairs in the groove. Electrostatic attraction stabilises the complex in a non-specific manner, while further stabilisation as well as specificity (recognition) is provided by hydrogen bonding and van der Waals interactions. The entire binding protein may contact the DNA on one side only, as in Cro (Anderson *et al*, 1981) and CAP (McKay and Steitz, 1981). Alternatively it may include a flexible region which wraps around the double stranded helix thus making contact on both sides, as in the case of Lambda repressor (Pabo *et al*, 1982) and RecA (Egelman and

Stasiak, 1986).

The general model appears to hold for many known examples, and homology among DNA binding proteins suggests even wider applicability (Sauer *et al*, 1982). While the model is directed towards an explanation of specificity, the elements it contains are fully compatible with non-specific binding. In this respect, the *N*-acylurea protein-DNA interaction, despite being based on synthetic protein constructs, might usefully be compared against it.

3.2 MATERIALS AND METHODS

3.2.1 *MATERIALS*

Highly polymerised calf thymus DNA (Product Code No. D-1501), deoxyribonuclease I (EC 3.1.21.1) (Product Code D-5025), ethidium bromide (Product Code E-8751) and heparin (Product Code H-3125) were obtained from the Sigma Chemical Co., St. Louis, Mo, U.S.A. Plasmid pBR322 DNA and endonuclease-free DNA polymerase I from *E. coli* (EC 2.7.7.7) were from Boehringer Mannheim GmbH, Penzberg, Western Germany. [Methyl,1′,2′-^{3}H] thymidine 5′-triphosphate (specific radioactivity 97Ci/mmol) was purchased from Amersham International, U.K. Sephadex G-50 was from Pharmacia Fine Chemicals, Uppsala, Sweden; nitrocellulose filter discs (Type HA, 25mm diameter, 0.45µm pore size) were from Millipore Corporation, Bedford, Mass, U.S.A., agarose was from BioRad Laboratories, Richmond, Ca, U.S.A., and scintillation fluid (type HP/b) was from Beckman-RIIC Ltd., High Wycombe, U.K. All other chemicals were of analytical grade.

3.2.2 *METHODS*

Shearing of Calf Thymus DNA

Calf thymus DNA approximately 12kb in size was prepared by passing a solution of DNA (0.1mg/ml) in 0.1SSC (SSC, 0.15M NaCl, 0.015M trisodium citrate) through a 25G-Yale syringe needle six times. Sheared DNA solutions were stored at -15°C.

Labelling of Calf Thymus and pBR322 DNAs

Sheared calf thymus DNA and the plasmid pBR322 were each labelled by the nick translation protocol (Rigby *et al*, 1977) in the presence of [methyl,1′,2′-^{3}H] thymidine 5′-triphosphate and subsequently purified by the spun column method using Sephadex G-50 (Maniatis *et al*, 1982). The products were stored at -15°C. Specific activities: 6.0 X 10^{6} - 1.0 X 10^{7} cpm/µg TCA-precipitated DNA.

Protein Determination

Protein was estimated quantitatively by the method of Lowry (1951) as mentioned in section 2.2.2.

Band Shift Assay

This assay, comprising agarose gel electrophoresis of protein-DNA and *N*-acylurea protein-DNA mixtures, was conducted as previously described (Huckett *et al*, 1986). Unlabelled pBR322 DNA was incubated with protein at 20°C for 30min as indicated in the legend to Figure 3.2, samples mixed with 1/3 volume of loading buffer (50%w/v sucrose, 4M urea, 0.1%w/v bromophenol blue, 50mM EDTA, pH 7.5) and analysed in 1.2%w/v agarose gels made and run with buffer containing 36mM Tris-HCl, 30mM sodium phosphate and 10mM EDTA (final pH 7.5). A BioRad Mini Sub DNA cell system was used; gel dimensions 100 X 65 X 3mm. Electrophoresis was at 40V for 3h. Gels were stained

after electrophoresis with ethidium bromide (EB) using conventional methodology (Maniatis *et al*, 1982), illuminated with transmitted UV light (300nm), and photographed using Kodak Tri-X Pan black and white film (ASA 400) in a single lens reflex camera fitted with a 55mm close-up lens and UV and red filters.

Nitrocellulose Filter Binding Assay

This assay was conducted essentially as previously described (Huckett *et al*, 1986) with modifications being made according to the requirements of individual experiments. All assays were carried out at 20°C.

For a straightforward assay, mixtures containing [^{3}H]calf thymus DNA or [^{3}H]pBR322 DNA and protein as indicated in the figure legends, in a final volume of 200µl assay buffer, were incubated for 30min, then filtered through nitrocellulose discs and rinsed with 2.0ml of the same buffer. The nitrocellulose filters were wetted in rinse buffer prior to use. Filtration and washing were conducted using a Millipore single glass filter holder under weak, constant Venturi suction. Finally, filters were air-dried and counted for radioactivity in scintillation fluid. In the initial stages of the work, the buffer used for incubation and washing was 100mM NaCl, 50mM Tris-HCl (pH 7.5), 5mM EDTA (assay buffer 1). Once it was realised that the binding reaction is sensitive to ionic strength, the buffer was changed to 50mM NaCl, 10mM Tris-HCl (pH 7.5), 1mM EDTA (assay buffer 2). For binding experiments involving insulin-albumin conjugates, the buffer was modified yet again to 50mM NaCl, 20mM Tris-HCl (pH 7.0), 1mM EDTA (assay buffer 3), in order to provide sufficient buffering at pH 7.0 to accomodate additions of conjugate in the form of stock solution at pH 4.0. Also, in some experiments, the DNA samples were isotope diluted. The variations employed for each particular experiment are made clear in the figure legends. In all other respects the ordinary assay procedure remained the same.

To investigate the effect of ionic strength on binding, protein-DNA complexes were allowed to form in the normal manner during a 30min incubation in assay buffer, then the incubation mixtures adjusted to Na^+ concentrations ranging from 25mM to 1.0M either by dilution with water or by the addition of 4M NaCl made up in appropriate buffer. Final total volume of each mixture after Na^+ adjustment was 200µl. Solutions were incubated for a further 10min before filtration. The filters were presoaked and subsequently rinsed with buffer which had been Na^+ adjusted to match the final incubation medium.

To investigate the effect of pH on binding, a series of special buffers was prepared for use with the assay which ranged from pH 4.0 to pH 10.0. In each buffer the total final concentration of Na^+ was adjusted to 80mM by the addition of NaCl as appropriate. Each buffer was used as incubation medium and washing medium in place of a regular assay buffer.

pH 4.0	50mM acetate buffer (sodium salt)
pH 5.0	50mM acetate buffer (sodium salt)
pH 6.0	50mM phosphate buffer (sodium salt)
pH 7.0	50mM phosphate buffer (sodium salt)
pH 8.0	50mM barbital buffer (sodium salt)
pH 9.0	50mM glycine-NaoH buffer
pH 10.0	50mM glycine-NaOH buffer

To investigate the temporal susceptibility of the binding interaction to anion competition, protein-DNA complexes were allowed to form in the normal manner in assay buffer, then, at a period of either 30sec or 30min after initial mixing, varying amounts of the polyanion heparin added in the form of freshly made up standard solution (0.1µg/µl in assay buffer) to give final total assay volumes of 200µl. Mixtures were incubated for a further 30sec before filtration.

3.3 RESULTS

3.3.1 *RELATIONSHIP BETWEEN CARBODIIMIDE MODIFICATION OF PROTEIN AND CAPACITY TO BIND DNA*

The entire series of *N*-acylurea CDI-albumins prepared at protein to carbodiimide mole ratios ranging from 1:10 to 1:2000 (section 2.2.2) was evaluated by means of the nitrocellulose filter assay for ability to bind both sheared calf thymus DNA and pBR322 DNA. Results, shown in Figure 3.1, demonstrate that while unmodified albumin and lightly carbodiimide-treated *N*-acylurea albumins could not bind DNA, *N*-acylurea albumins prepared at protein to carbodiimide mole ratios of 1:250 and greater were able to bind DNA. Albumins in this category had already been shown to carry more than 13 *N*-acylurea moieties per albumin molecule (Table 2.1) and to exhibit significant electropositivity as a result (Figure 2.3). The first *N*-acylurea albumin in the mole ratio series which significantly bound both calf thymus (linear) and plasmid (circular) DNA was the 1:500 derivative (Figure 3.1), previously shown to have *N*-acylurea groupings at 27 sites per molecule (Table 2.1). On the basis that this derivative appeared to provide reliable DNA binding with least modification, *N*-acylurea albumins prepared at protein to carbodiimide mole ratios of 1:500 were deemed most suitable for conjugation to insulin.

The capacity of proteins to bind DNA once they have been converted to *N*-acylurea derivatives is illustrated by alternative means in Figure 3.2, which shows the result of DNA gel electrophoresis following incubation of pBR322 DNA samples with varying amounts of proteins. Retardation of migration resulting from the complexing of plasmid to protein (the band shift effect) was apparent only in the case of *N*-acylurea protein incubations. Band shift analysis of albumin and *N*-acylurea CDI-albumin (1:500) is presented in Figure 3.2A. In the presence of 0.33µg pBR322 DNA, 4µg A-CDI produced a distinct

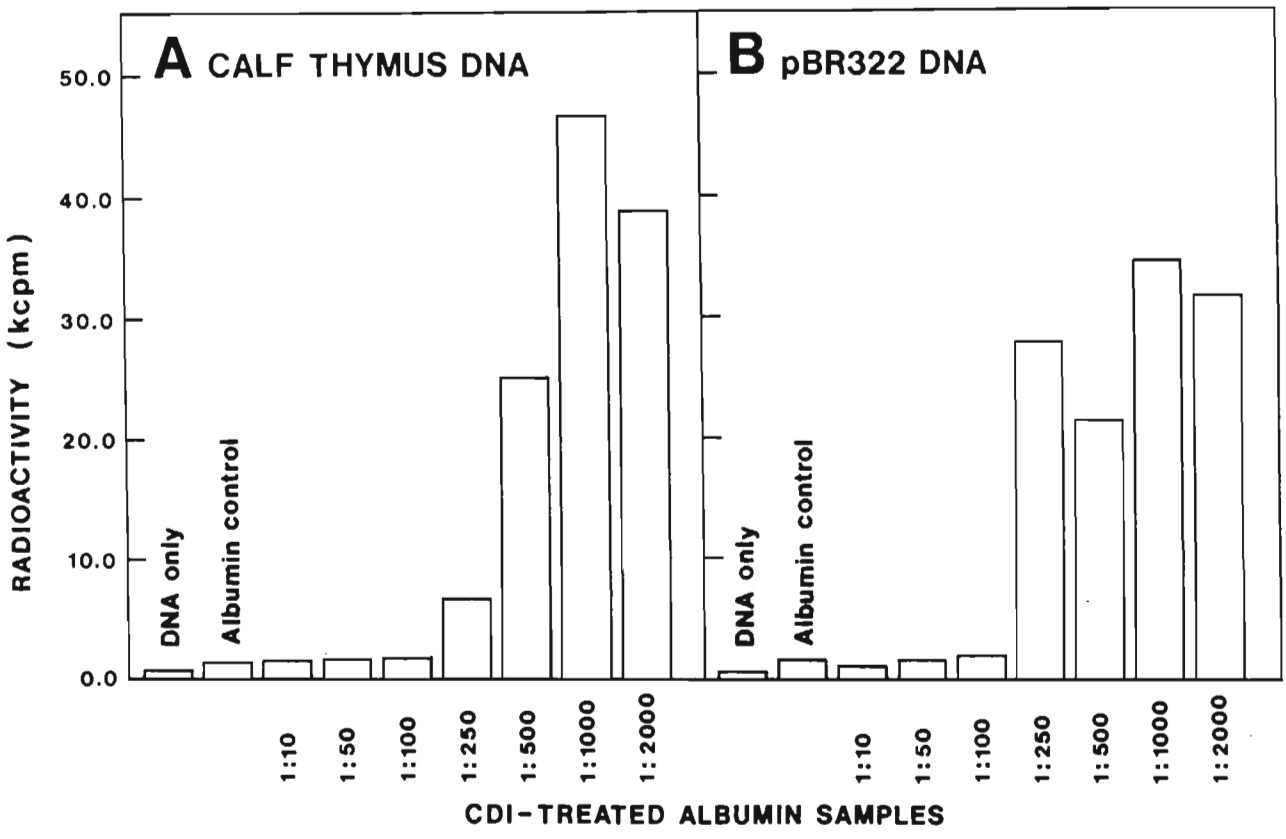


Figure 3.1 Binding of DNA to *N*-acylurea albumins prepared with *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride at protein to carbodiimide mole ratios varying from 1:10 to 1:2000

DNA binding was assayed by the nitrocellulose filter technique using buffer 1 (high ionic strength). Each assay mixture contained 10ng DNA and 8.8µg protein.

A. [^{3}H]calf thymus DNA (6.5 X 10^4 cpm per assay) with the complete range of *N*-acylurea albumins.

B. [^{3}H]pBR322 DNA (7.6 X 10^4 cpm per assay) with the complete range of *N*-acylurea albumins.

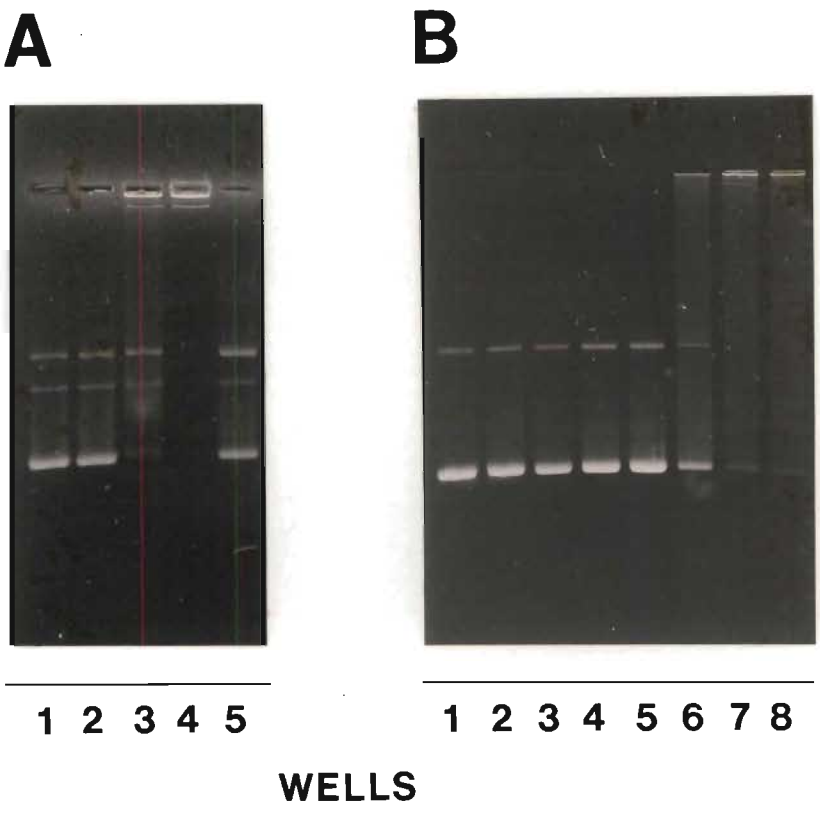


Figure 3.2 Band shift assay of binding interactions between proteins and DNA

Samples of pBR322 DNA were incubated with protein at 20°C for 30min prior to agarose gel electrophoresis. Each incubation mixture contained pBR322 DNA and protein as indicated, in a final volume of 10µl buffer (50mM NaCl, 20mM Tris-HCl (pH 7.0), 1mM EDTA).

A. pBR322 DNA, 0.33µg per well. 1, pBR322, no protein; 2, pBR322 plus albumin (4.0µg); 3 and 4, pBR322 plus A-CDI (4.0µg and 8.0µg respectively); 5, pBR322 plus cytochrome C (4.0µg).

B. pBR322, 0.45µg per well. 1, pBR322, no protein; 2 and 3, pBR322 plus I-A (2.0µg and 3.3µg respectively); 4-8, pBR322 plus I-[A-CDI] (0.4µg, 1.0µg, 1.66µg, 2.66µg and 3.3µg respectively).

Conjugates were added in the form of purified stock solutions (pooled Sephadex G-100 fractions, as shown in Figure 3.3).

band shift, particularly in relation to the supercoiled form of the plasmid, while at 8µg A-CDI the plasmid was fully bound and the resultant complex virtually unable to enter the gel. In contrast, unmodified albumin and the basic protein cytochrome C were without effect on pBR322 DNA migration under the same electrophoretic conditions. A comparable analysis of insulin-albumin conjugates is presented in Figure 3.2B. In the presence of 0.45µg pBR322 DNA, band shift became observable at 1.0µg I-[A-CDI] among the supercoiled plasmid population, and at 1.66µg I-[A-CDI] significant retardation of both supercoiled and nicked circular forms was seen as continuous smudging of the DNA, while at higher levels of the same conjugate the complex was quite unable to migrate. I-[A-CDI], being particularly large and yet heterogeneous in size, appeared to bring about dramatic band-shifts of variable dimensions, thus giving rise to more "blurring" than did A-CDI. Like unmodified albumin, I-A was without effect on pBR322 mobility. Insulin was similarly without effect (result not shown).

3.3.2 *DNA BINDING CAPACITY OF INSULIN-ALBUMIN CONJUGATES FRACTIONATED BY GEL FILTRATION*

Purified conjugates obtained by gel filtration through Sephadex G-100 (Figure 2.6) were assayed for DNA binding activity by the nitrocellulose filtration technique. Figure 3.3 illustrates typical binding interactions of pBR322 DNA with the various size species of I-A and I-[A-CDI] obtained by fractionation during the purification procedure. Each assay mixture contained 1.5µg conjugate. At this level of protein the DNA present in the assay mixture (10ng) is not limiting (Figure 3.4B). The capacity to bind DNA was demonstrated by I-[A-CDI] but not by I-A. While all fractions of I-[A-CDI] underwent significant binding, binding was lowest for the very large polymers and there was a reproducible trend for the binding capacity of the larger species of conjugate to be inversely proportional to molecular size (fractions 15 - 20). This is likely to be a

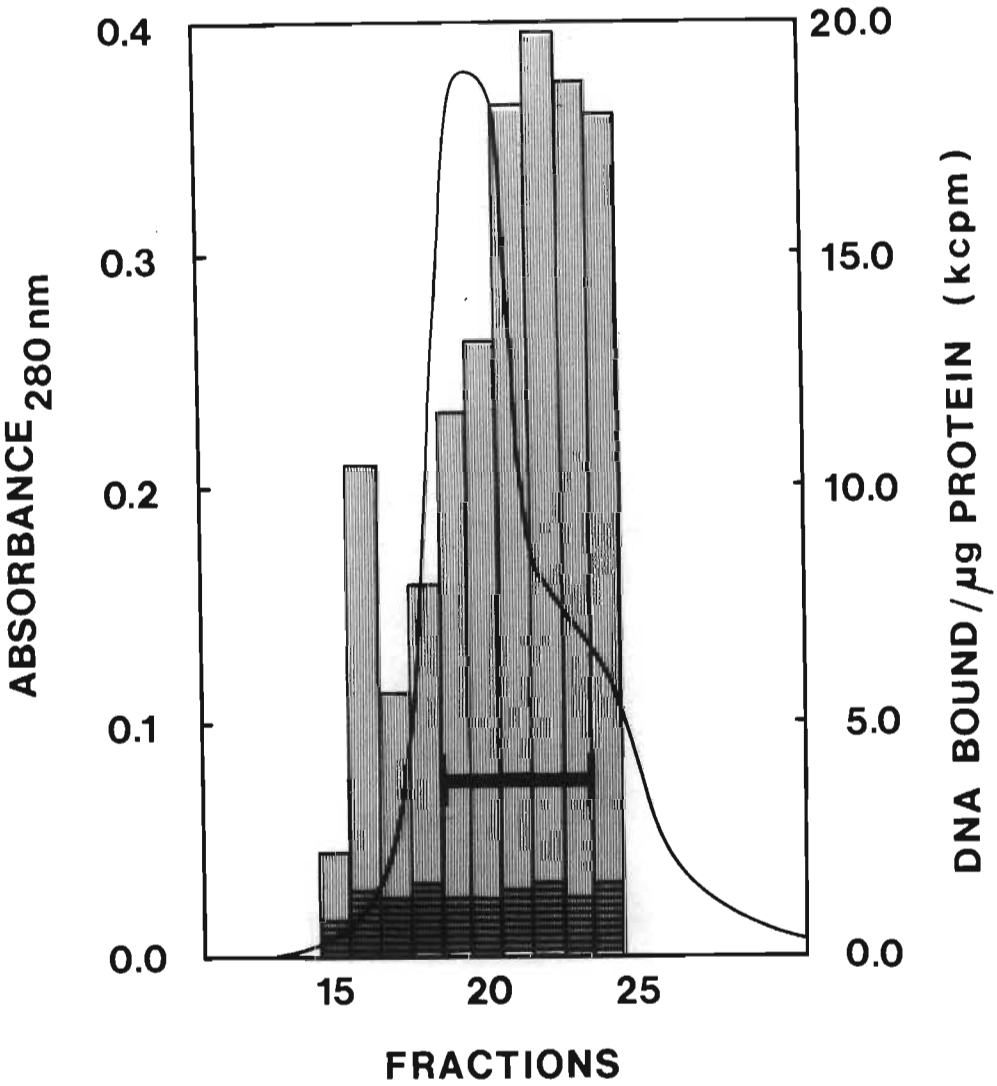


Figure 3.3 DNA binding capacity of insulin-albumin conjugates fractionated by gel filtration

DNA binding was assayed by the nitrocellulose filtration technique using assay buffer 3. Each assay mixture contained 4.5ng [^{3}H]pBR322 DNA (3.0 X 10^4 cpm), 5.5ng unlabelled pBR322 DNA and 1.5µg protein conjugate.

Absorption at 280nm of I-A and I-[A-CDI] (——); DNA bound by I-[A-CDI] (light hatching); DNA bound by I-A (heavy hatching).

The horizontal bar indicates those fractions pooled for use in binding and transfection experiments.

reflection of the reduction in the ratio of surface area to weight that accompanies polymerisation. Maximal binding was achieved and maintained over a range of the more moderately sized species (fractions 21 - 24).

On the basis of (i) DNA binding capacity (Figure 3.3), (ii) insulin content (Figure 2.8) and (iii) concentration (Figures 2.6 and 2.8), I-[A-CDI] fractions 19 to 23 inclusive were pooled to provide an I-[A-CDI] stock solution for use in all experiments requiring purified conjugate. Fractions of I-A were pooled in an equivalent manner. Stock solution concentrations ranged from 0.38 to 0.42µg protein/ml.

3.3.3 *PARAMETERS OF BINDING*

The Effect of Protein Concentration

The interactions of DNA with albumins and insulin-albumin conjugates at varying concentrations of protein are shown in Figure 3.4. Results confirm that only *N*-acylurea derivatives undergo binding. Total DNA per reaction mixture was 10ng in all cases. The amount of DNA bound increased linearly with increasing protein until a point of saturation (equilibrium) was reached; beyond this point no more DNA could be taken up into the protein-DNA complex.

Figure 3.4A illustrates the binding of [^{3}H]calf thymus DNA (6.5 X 10^4 cpm) and [^{3}H]pBR322 DNA (7.6 X 10^4 cpm) to *N*-acylurea CDI-albumin and *N*-acylurea Me^+CDI-albumin respectively, each of the two proteins having been carbodiimide-modified at a 1:500 mole ratio. Results indicate that A-Me^+CDI, which was fully saturated with DNA at an input of 0.05µg protein, was able to enter into a more tenaceous binding relationship than A-CDI, 0.10µg of which had to be present for saturation point to be reached. Thus the observed equilibrium constant (K_{obs}) was greater for A-Me^+CDI than for A-CDI. This result is in keeping

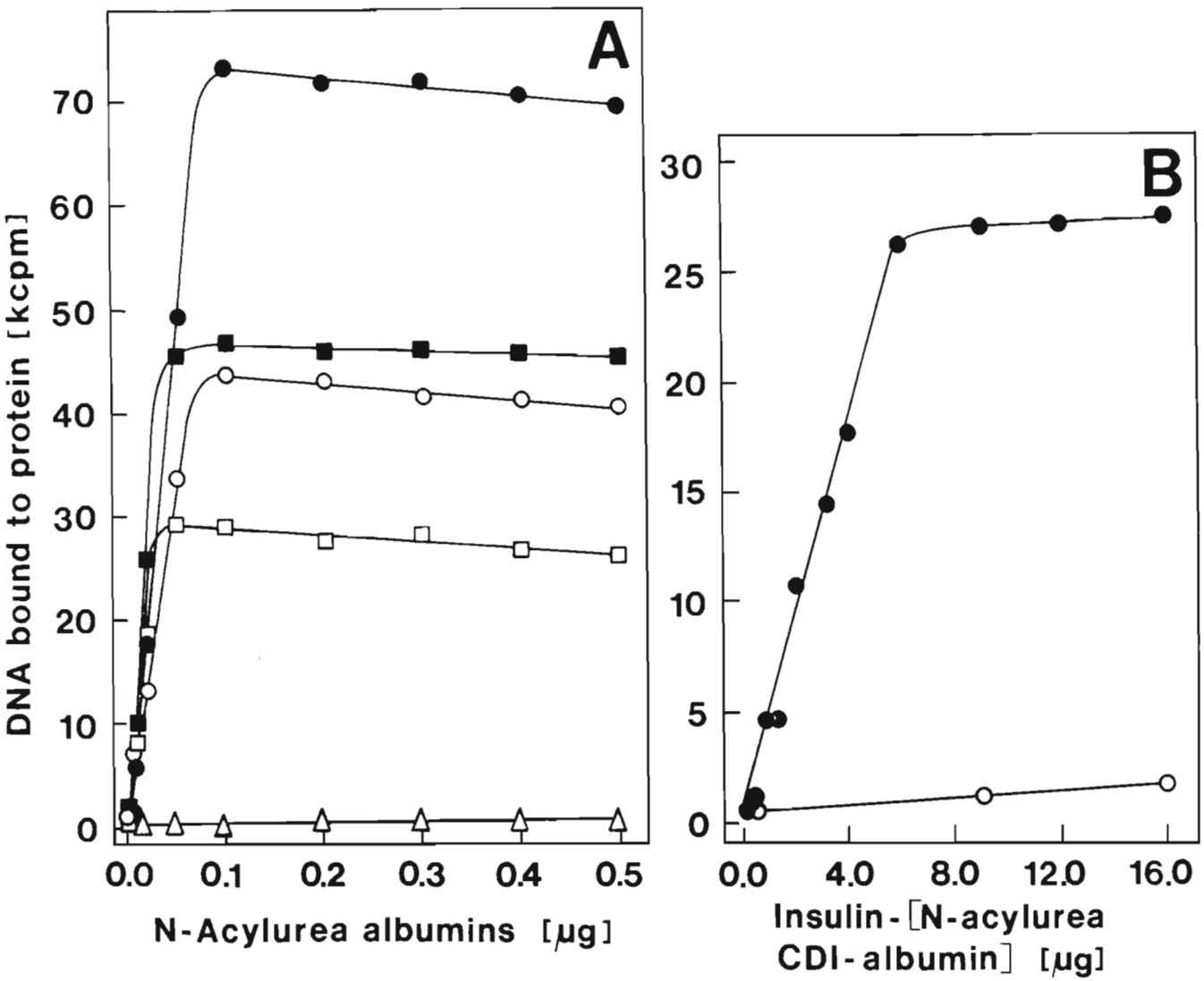


Figure 3.4 Binding of DNA to *N*-acylurea proteins at varying protein concentration

DNA binding was assayed by the nitrocellulose filtration method.

A. Each assay mixture contained 10ng [^{3}H]pBR322 DNA (7.6 X 10^4 cpm) or 10ng [^{3}H]calf thymus DNA (6.5 X 10^4 cpm) and protein as indicated, in assay buffer 2. Modification of albumin was at a protein to carbodiimide mole ratio of 1:500. A-CDI with pBR322 (●) and calf thymus DNA (○); A-Me$^+$CDI with pBR322 (■) and calf thymus DNA (□); albumin control with either DNA (△).

B. Each assay mixture contained 4.5ng [^{3}H]pBR322 DNA (4.5 X 10^4 cpm) and 5.5ng unlabelled pBR322 DNA together with protein as indicated, in assay buffer 3. Each protein conjugate was added as purified stock solution (Figure 3.3). I-[A-CDI], (●); I-A control (○).

with the expectation that the quaternary nitrogen functions present in the *N*-acylurea moieties of A-Me$^+$CDI would give rise to a higher affinity for DNA than exhibited by A-CDI, which is characterised by tertiary nitrogens. That A-Me$^+$CDI is more positively charged than its A-CDI counterpart has already been noted (Figure 2.3A). However, it is interesting to observe that in spite of its apparently higher affinity for DNA, A-Me$^+$CDI showed a lower overall DNA binding capacity: at saturation it carried less DNA than did A-CDI. The explanation for this is not immediately clear, but could be related to a reduction in the number of available binding sites as a result of conformational constraints.

Results shown in Figure 3.4A also indicate that pBR322 DNA was bound more effectively than sheared calf thymus DNA by both *N*-acylurea albumins. This could be explained simply by the loss of particularly small or large fragments of calf thymus DNA during the filtration and washing procedures of the assay. Alternatively, it could be explained by a higher *N*-acylurea protein binding affinity for circular (particularly supercoiled) forms of DNA, or a more efficient spatial accomodation of this form of DNA on the protein surface, or both. The results of band shift assay (Figure 3.2) indicate that from a mixture of supercoiled and nicked circular forms of plasmid, the former was indeed selectively bound by *N*-acylurea proteins.

Of interest in relation to both of the above sets of findings is the observation that the interaction of pBR322 DNA with A-CDI was the only one leading to virtual exhaustion of free DNA in the incubation mixture (7.2×10^4 cpm bound).

Results presented in Figure 3.4B illustrate the binding of isotope diluted [^{3}H]pBR322 DNA (4.5×10^4 cpm) to insulin-[*N*-acylurea CDI-albumin]. Saturation point was reached at an input of 8.0µg conjugate. This is 80 times the amount of unconjugated A-CDI required for saturation under the same conditions. Since I-[A-CDI] is not larger than A-CDI by as

much as a factor of 80 (section 2.4), K_{obs} for the conjugate is significantly lower than for both of the unconjugated *N*-acylurea albumins. The lower value is compatible with the fact that the components of the conjugate lose some electropositivity during the glutaraldehyde coupling procedure such that the final product is only moderately basic (Figure 2.3B). Nevertheless the capacity of the conjugate to bind DNA is significant.

The Effect of NaCl Concentration

The dissociation by Na^+ of *N*-acylurea protein-DNA complexes is shown in Figure 3.5. Figure 3.5A illustrates the salt induced dissociations of *N*-acylurea CDI-albumin and *N*-acylurea Me^+CDI-albumin from [^{3}H]pBR322 DNA, while Figure 3.5B illustrates the comparable dissociation of the conjugated binding protein insulin-[*N*-acylurea CDI-albumin] from [^{3}H]pBR322 DNA.

As deduced from the curves, concentrations of NaCl required for the half-dissociation of individual *N*-acylurea protein-DNA complexes are: A-CDI, 0.15M; A-Me^+CDI, 0.24M; I-[A-CDI], 0.1M. These values confirm previous observations (Figure 3.4) that the affinity for DNA (K_{obs}) of A-Me^+CDI is higher than that of A-CDI, and that the affinity of the I-[A-CDI] conjugate for DNA is weaker than the DNA binding affinities of either of the unconjugated *N*-acylurea proteins. Moreover these values are fully correlative with observed charge properties (Figure 2.3).

That the sodium ion caused almost complete dissociation of the complexes formed between unconjugated *N*-acylurea albumins and DNA (Figure 3.5A) suggests an important role for electrostatic interaction in the binding phenomenon. Of considerable interest and contrast in this respect is the residual binding displayed by I-[A-CDI]-DNA complexes at high concentrations of sodium (Figure 3.5B), which possibly indicates the involvement of strong additional forces in the binding reaction between conjugate and DNA. In practical terms the results pinpoint the

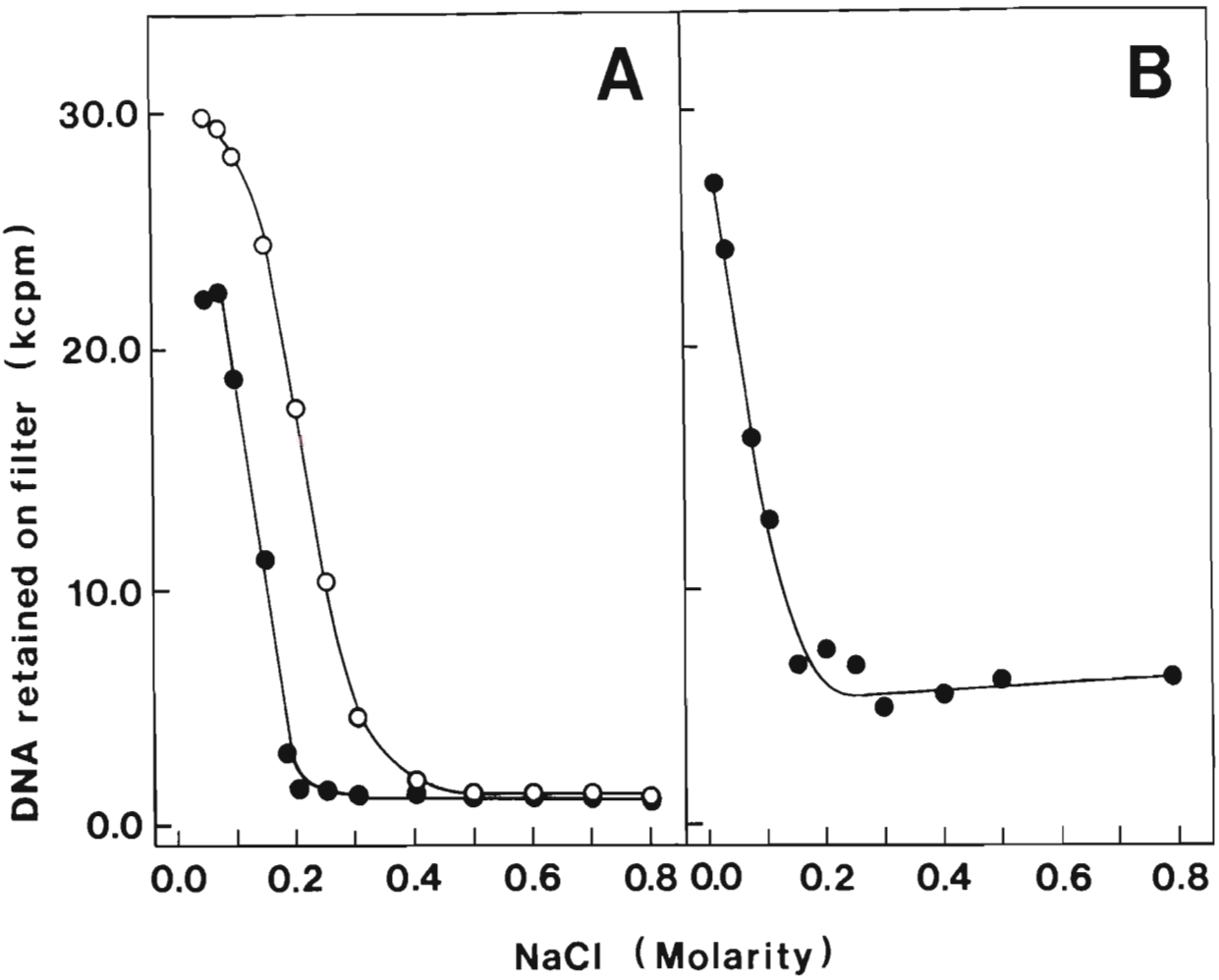


Figure 3.5 The effect of NaCl concentration on the dissociation of *N*-acylurea protein-DNA complexes

DNA binding was assayed by the nitrocellulose filtration method.

A. Each assay mixture contained 5ng [^{3}H]pBR322 DNA (3.8 X 10^4 cpm) together with 100ng A-CDI or 80ng A-Me$^+$CDI in assay buffer 2. Both *N*-acylurea albumins were prepared at a protein to carbodiimide mole ratio of 1:500. Dissociation of pBR322 from A-CDI (●); dissociation of pBR322 from A-Me$^+$CDI (○).

B. Each assay mixture contained 4.5ng [^{3}H]pBR322 DNA (4.1 X 10^4 cpm) and 5.5ng unlabelled pBR322 DNA together with 5.5µg I-[A-CDI] (purified stock solution) in assay buffer 3.

importance of using a low-salt medium (assay buffers 2 and 3) for *in vitro* binding studies, and emphasise ionic strength as a factor for serious consideration in future transfection work whether with cells *in vitro* or whole organisms *in vivo*.

The Effect of pH

Conjugated proteins were used as DNA binding models to study the effect of pH on the binding reaction. Figure 3.6 shows the amount of [^{3}H]pBR322 DNA bound by 4.0µg insulin-albumin (control) and 4.0µg insulin-[*N*-acylurea CDI-albumin] at pH values ranging from 4.0 to 10.0. In order to span this pH range, four different types of buffer were necessarily employed (section 3.2.2); therefore salts of sodium were selected for use in each buffer system and the final sodium concentration adjusted to 80mM in every case so as to eliminate variations in binding due to the influence of ionic strength and Na^+ concentration in particular.

Results showed that

(i) unmodified I-A bound DNA at low pH (pH 4.0 - 5.0);
(ii) I-[A-CDI] lost its capacity to bind DNA at high pH (pH 8.0 - 10.0), and
(iii) optimal conditions for controlled nitrocellulose filter binding assays existed only close to neutral pH (between pH 6.0 and 7.0) .

These results demonstrate the strong pH dependence of the protein-DNA binding reaction as determined by the nitrocellulose filter assay. They can be explained strictly in terms of the protonation and dissociation of side chain amino and *N*-acylurea groupings, and in this respect they fully support the idea that *N*-acylurea protein-DNA binding is charge dependent and at least partially electrostatic in nature. However, it is not known to what extent pH differences affect the protein binding properties of nitrocellulose. Explanations which account only for charge alterations on the proteins are likely, therefore, to be

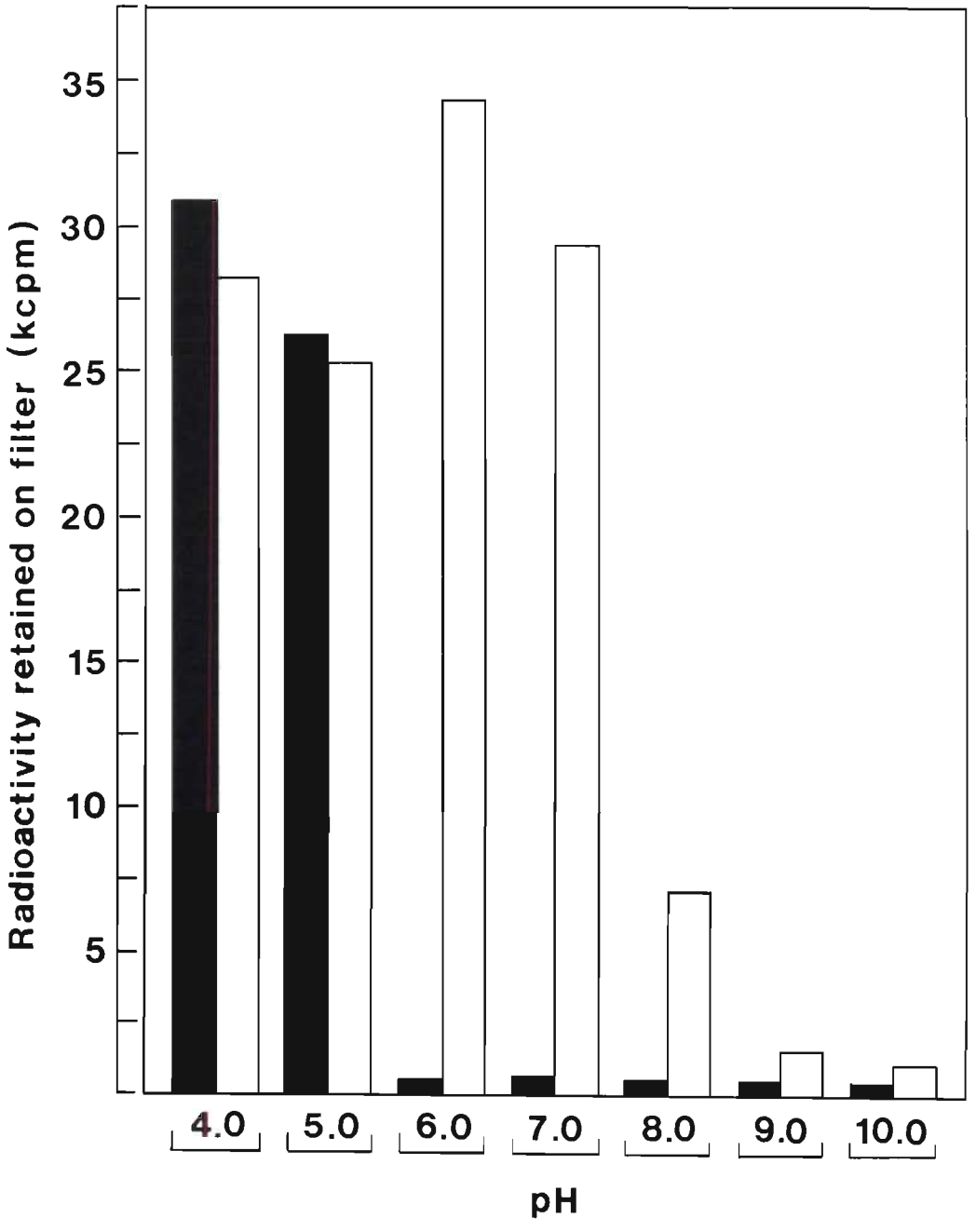


Figure 3.6 pH Dependency of the binding of DNA to insulin-albumin and insulin-[*N*-acylurea CDI-albumin] assayed by the nitrocellulose filtration technique

Each assay mixture contained 4.5ng [^{3}H]pBR322 DNA (4.2 X 10^4 cpm), 5.5ng unlabelled pBR322 DNA, and 6.0µg purified I-A (■) or I-[A-CDI] (□). Both incubation and washing steps were conducted using one of a series of buffers as defined in Materials and Methods.

simplistic.

One clearly important result which does emerge is that the nitrocellulose filter binding assay must be conducted under carefully controlled conditions of pH close to neutral. This conclusion reinforces the importance of having redesigned the nitrocellulose filter assay buffering medium for use with insulin-albumin conjugates (assay buffer 3); it was changed to include a slightly higher concentration of Tris-HCl in order to accomodate additions of purified conjugate in low pH acetate buffer without the final assay pH going beyond acceptable bounds.

The Effect of Time

Table 3.1 demonstrates that when *N*-acylurea protein and DNA are brought together in solution, the formation of filter-detectable complex is rapid, occurring in less than 30 seconds, the complex remaining stable for at least 50 minutes thereafter.

Table 3.1 Effect of time of incubation on *N*-acylurea protein-DNA binding determined by the nitrocellulose filtration assay

Each assay mixture contained 10ng [^{3}H]calf thymus DNA (6.5×10^{4} cpm) and 1.0µg A-CDI (1:500) in assay buffer 2.

TIME OF INCUBATION (min)	DNA BOUND TO PROTEIN (kcpm)
0.5	29.1
12.5	29.0
25.0	26.5
37.5	24.0
50.0	28.8
50.0 (no protein)	0.7

3.3.4 *HEPARIN CHALLENGE EXPERIMENTS: THE BINDING PROCESS*

Following initial reaction between *N*-acylurea proteins and DNA, the complexes formed were challenged with heparin at two different periods of elapsed time and the residual binding measured by nitrocellulose filter assay. Figure 3.7 shows the result of heparin competition on the complexes formed between *N*-acylurea CDI-albumin and calf thymus DNA (A), *N*-acylurea CDI-albumin and pBR322 DNA (B), and insulin-[*N*-acylurea CDI-albumin] and pBR322 DNA (C). In their initial state, 30 seconds after mixing, all complexes formed appeared to be susceptible to dissociation when challenged by heparin, suggesting a simple, rapidly formed, ion-dependent (electrostatic) binding reaction. However, after allowing the reaction to proceed for 30 minutes, heparin challenge was significantly less effective, indicating the development over that period of time of tighter binding involving at least one non-electrostatic component.

3.3.5 *FURTHER INVESTIGATIONS INTO THE NATURE OF BINDING*

Minor experiments directed towards obtaining additional information on the nature of the binding interactions between *N*-acylurea proteins and DNA were undertaken. Heating *N*-acylurea CDI-albumin-pBR322 complexes to 55°C for 10min had no effect on binding. Treating an identical complex with phenol at 20°C did not lead to separation of the plasmid into the aqueous phase; however, when the phenol contained 0.1%SDS and the extraction was at 60°C for 5 min, the plasmid was recovered in the aqueous phase and was shown by gel electrophoresis to be mainly in the supercoiled form. These results indicate that the complex of *N*-acylurea CDI-albumin and pBR322 DNA is resistant to moderate temperatures, is held together by noncovalent bonds, and does not require or cause nicking of the DNA.

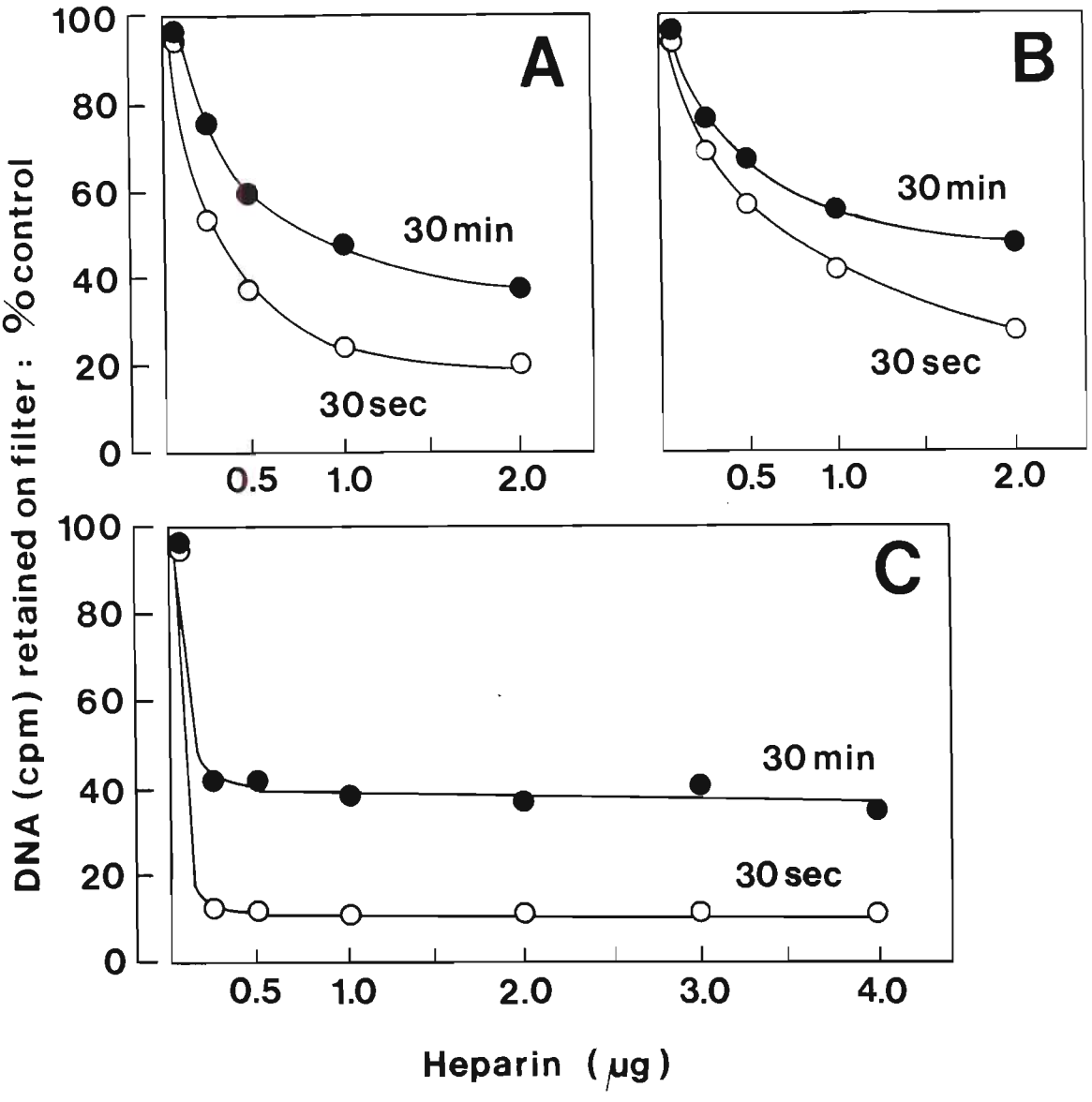


Figure 3.7 Time dependency of heparin challenge on *N*-acylurea protein-DNA binding assayed by the nitrocellulose filtration method

Assay mixtures were incubated for either 30sec or 30min prior to addition of heparin, then incubation continued for a further 30sec before filtration.

A. Heparin challenge to the binding of calf thymus DNA to *N*-acylurea CDI-albumin (1:500). Each assay mixture contained 10ng [^{3}H]calf thymus DNA (6.5 X 10^4 cpm) and 1.0μg A-CDI in assay buffer 2.

B. Heparin challenge to the binding of pBR322 DNA to *N*-acylurea CDI-albumin (1:500). Each assay mixture contained 10ng [^{3}H]pBR322 DNA (7.6 X 10^4 cpm) and 1.0μg A-CDI in assay buffer 2.

C. Heparin challenge to the binding of pBR322 DNA to insulin-[*N*-acylurea CDI-albumin]. Each assay mixture contained 4.5ng [^{3}H]pBR322 DNA (3.2 X 10^4 cpm), 5.5ng unlabelled pBR322 DNA and 6.2μg I-[A-CDI] in assay buffer 3.

3.4 SUMMARY AND DISCUSSION

Using both sheared calf thymus DNA (linear, heterogeneous) and pBR322 DNA (circular, homogeneous), the DNA binding reactions of various *N*-acylurea albumins and an insulin-conjugated derivative have been investigated. Two methods were employed: the agarose gel electrophoresis band shift assay, and the nitrocellulose filter assay. Both of these assays were used at the outset to establish the ability of *N*-acylurea albumins to bind DNA and to demonstrate the fundamental dependence of the binding reaction on the presence of *N*-acylurea carbodiimide adducts (Figures 3.1 and 3.2). In addition, the precision of the nitrocellulose filter technique allowed a comparative assessment to be made of the binding capacities of the complete series of CDI-treated albumins, resulting in identification of the 1:500 mole ratio products as most suitable for conjugation to insulin on grounds of significant ability to bind different forms of DNA (Figure 3.1). Analysis of the DNA binding capacity of gel filtration fractions of insulin-[*N*-acylurea CDI-albumin] after post-conjugation purification was also made possible by the nitrocellulose filtration technique (Figure 3.3).

The nitrocellulose filter method was utilised further to investigate various aspects of the DNA binding reaction (Figures 3.4 - 3.7). For this purpose, *N*-acylurea CDI-albumin (1:500) and *N*-acylurea Me^{+}CDI-albumin (1:500) were chosen as important study models, being the modified albumins carrying different *N*-acylurea carbodiimide moieties deemed suitable for conjugation to insulin. Insulin-[*N*-acylurea CDI-albumin] was necessarily the conjugated *N*-acylurea protein selected for study, being the only one destined for possible future use in transfection work on the basis of solubility (section 2.4). By means of the experiments undertaken with these *N*-acylurea proteins it was possible to

- demonstrate the effect of *N*-acylurea protein concentration on its binding to DNA, and from the binding curves determine the relative affinities for DNA of the different *N*-acylurea proteins;

- show that *N*-acylurea protein-DNA complexes are sensitive to NaCl concentration, and from the dissociation curves determine and confirm relative affinities for DNA of the different *N*-acylurea proteins;

- show that the binding reaction is pH dependent;

- show that the binding reaction occurs very rapidly, and

- deduce the existence of more than one mode of interaction between *N*-acylurea proteins and DNA.

Some of the binding studies using *N*-acylurea albumin have been reported previously in conjunction with supportive observations made using *N*-acylurea transferrins (Huckett *et al*, 1986).

A number of factors indicate that there is a strong electrostatic component in the binding interaction:

- *N*-acylurea Me^{+}CDI-albumin, which is more strongly basic than the eqivalent *N*-acylurea CDI-albumin (Figure 2.3A), shows an apparently higher equilibrium constant (K_{obs}) (Figures 3.4A and 3.5A); likewise, the insulin-[*N*-acylurea CDI-albumin] conjugate, which is less basic than either of the two *N*-acylurea albumins (Figure 2.3B), has a correspondingly lower K_{obs} (Figures 3.4B and 3.5B);

- K_{obs} for all *N*-acylurea protein-DNA interactions are dependent on monovalent ion concentration ($[Na^{+}]$ in the experiments reported here); K_{obs} is reduced with increasing $[Na^{+}]$ (Figure 3.5), and

- the polyanion heparin is able to compete with DNA for protein binding sites thus reducing K_{obs} (Figure 3.7).

Using thermodynamic analysis as developed by Record *et al* (1976) for the binding of charged ligands to nucleic acids, estimates have been made of the number of charge-charge interactions occurring in various *N*-acylurea protein-DNA complexes (Ariatti and Hawtrey, 1987); they appear to be of the same order as those calculated by Record and colleagues for several naturally occurring DNA binding proteins. While the role of electrostatic interaction at the binding site is thus strongly supported, and might be a dominant one, additional non-electrostatic forces do appear to be operative in the binding reaction as indicated by

- the inability of proteins such as cytochrome C, of similar basicity to *N*-acylurea albumins (Figure 2.3), to bind DNA, and

- the existence of at least two modes of binding interaction, one of which is not easily disrupted by anionic competition, as deduced from heparin challenge experiments (Figure 3.7).

Evidence for non-electrostatic interaction between *N*-acylurea entities and nucleic acid bases, particularly unpaired guanine residues, has been provided by H^1Nmr spectrometry which demonstrates a downfield shift in the signal from the amino group of guanosine in solution in DMSO in the presence of *N*-acetylurea (Ariatti and Hawtrey, 1987). The observed shift is explicable by hydrogen bonding; model building was used by the same authors to show that three hydrogen bonds are possible with guanine whereas at best only two bonds are possible with A, C, U or T. An alternative explanation is a glyoxal-type addition reaction, also shown to be feasible by modelling. Such interactions imply binding within single-stranded regions of DNA. The preferential binding of *N*-acylurea CDI-albumin and its conjugate for supercoiled DNA, noted in Figure 3.2 and also

reported previously with *N*-acylurea CDI-transferrin (Huckett *et al*, 1986) and *N*-acylurea Me^{+}CDI-transferrin (Ariatti and Hawtrey, 1987), does in fact support this possibility, since DNA in the supercoiled form contains more single-stranded regions than relaxed closed circular or linearised DNA configurations.

The heparin challenge experiments (Figure 3.7) indicate the existence of two modes of binding interaction which cannot be discriminated by nitrocellulose filtration assay (Table 3.1), both of which are destroyed by salt (Figure 3.5): (i) fast-forming heparin sensitive bonds which are presumably mainly electrostatic, and (ii) interactive forces which come into play more slowly, resulting in greater binding stability and resistance to anionic competition, which probably involve non-electrostatic mechanisms. The observed salt sensitivity indicates that the latter interaction requires the electrostatic component for full stabilisation. In interpreting results of heparin challenge on the binding of RNA polymerase to restriction fragments of T7D111 DNA, Melancon and coworkers (1982) deduced the existence of at least three categories of binding interaction, including the two noted here, and showed that they occurred on certain subsets of fragments and in certain positions therein. In the case of the *N*-acylurea protein-DNA interaction, the two modes of binding might be interpreted as reflecting either (i) two different categories of DNA sequence giving rise to two different types of binding site, one of which could be more specific and more stable than the other, or (ii) two different types of interaction at each of many non-specific binding sites at which DNA sequence is of little importance.

In the light of both the results obtained and known models of nucleoprotein complex formation, it is suggested that the binding interaction consists of the following: a degree of close physical contact occurs between the *N*-acylurea protein and the DNA which is very rapidly stabilised in a non-specific manner by electrostatic forces in the form of ion pairing between

N-acylurea entities and nucleic acid phosphates. Localised unwinding of the DNA double helix might be induced at this stage by unknown mechanisms. Thereafter the complex is stabilised further in a slower process of accomodation leading to hydrogen bonding and possibly other additional interactions. This latter process would be somewhat site-dependent, although not site-specific, and therefore it is envisaged that accomodation might include a degree of movement between the surface of the protein and the phosphodiester backbone of the DNA prior to the bonding taking place. A precedent exists for this idea: in proposing a model for the non-specific binding shown by CAP, Weber and Steitz (1984) have suggested that the protein is capable of sliding in and out of the major grooves of the DNA to sample the sequence until complementarity of structure leading to hydrogen bonding is achieved. Those authors conceive that electrostatic interaction provides initial long range energy for the formation of the complex as well as the orientation of the protein and DNA entities in relation to one another, while hydrogen and other types of bonds represent short range binding energies which come into play secondarily. They base their model on thermodynamic calculations which suggest that electrostatic stability is maximum at a protein-DNA separation distance of 12 Å at which hydrogen bonds or other binding contacts of a directional nature do not form. Charge-charge interactions are clearly not merely binding initiators, however. Evidence from Na^+ dissociation studies (Figure 3.5; Melancon *et al*, 1982) indicate that electrostatic interactions continue to play an important role in complex stabilisation even after secondary bonding has taken place as demonstrated by heparin competition (Figure 3.7).

The residual binding displayed by insulin-[*N*-acylurea CDI-albumin] conjugate at high [Na^+] (Figure 3.5B), not seen in the dissociation of *N*-acylurea albumins (Figure 3.5A), is of particular interest in view of the possible role of the conjugate in gene transfer. In terms of the model described above, a residual, Na^+-resistant bond would be likely to

comprise one of the non-electrostatic modes of interaction. One possibility is that sufficient hydrogen bonding had taken place between the I-[A-CDI] conjugate and DNA that electrostatic interaction was not required to stabilise and maintain overall binding free energy, although in view of the salt sensitivity of other complexes, including tightly bound examples (Melancon *et al*, 1982), this seems unlikely. Another possibility is that a type of bonding such as the formation of glyoxal-type adducts as suggested by Ariatti and Hawtrey (1987) had taken place; bonding of this type would certainly not be sensitive to ionic competition, although it is not known whether it could maintain the integrity of the complex without the contribution of charge-charge interactions. Both of the aforementioned suggestions are based on an assumption of complete solubility of binding components and bound complex. A further possibility is that a proportion of the I-[A-CDI]-DNA assay mixture had coprecipitated out of solution in the dissociation experiments. While no visual evidence of precipitation was found, this possibility was taken seriously because of the insolubility problems encountered with insulin-[*N*-acylurea Me^{+}CDI-albumin]-DNA complexes and the vital importance of conjugate-DNA complex solubility to the design of future gene transfer experiments. Interestingly, in preparing protein-nucleic acid conjugates for use as immunogens, Stollar (1980) has noted that in certain circumstances carbodiimide-modified proteins may form precipitates with DNA. In the present work, therefore, appropriate precautions were taken in the transfection protocol to deal with the possibility of insoluble complex formation (Chapter 5).

CHAPTER FOUR

BINDING OF INSULIN AND INSULIN CONJUGATES TO HepG2 CELLS

4.1 INTRODUCTION

Techniques designed to measure the binding of a ligand to its receptor are fundamental tools in a number of different research disciplines, since effector molecules as diverse as neurotransmitters, antibodies, lectins, drugs and hormones all act by triggering a cascade of biochemical reactions through the initial formation of a complex with the cognate receptor situated in the cell surface plasma membrane or in one of the organellar membrane systems. In almost all cases, detection and measurement of the binding interaction requires a sensitive method: most natural ligands occur at extremely low physiological concentrations, and their receptors exist in femtomole to picomole quantities per milligram of membrane protein, which means a receptor concentration of 1×10^{-12} - 1×10^{-8}M in a binding assay (Levitzki, 1985). The detection methods first devised and still usually employed are forms of radioligand assay, in which the ligand is labelled to high specific radioactivity using a suitable isotope. The one of choice is [^{125}I], not only because a product of high specific radioactivity (150 -1500 kCi/mol) is obtained, but also because the half life ($t_{1/2}$ = 60.2 days) is moderately long and the radioiodination procedure easy to perform in the laboratory (Gammeltoft, 1984). Of the various methods available for the iodination of peptides and proteins at the time when ligand-receptor binding studies were being initiated in the late 1960s and early 1970s, the one most commonly used was that of Hunter and Greenwood (1962) involving chloramine-T oxidation of sodium iodide to the cation I^+ conducted under buffered aqueous

conditions in the presence of the protein substrate. The I^+ becomes incorporated into the ring structure of the amino acid tyrosine directly and rapidly in a single stage reaction at room temperature. The iodinated protein may easily then be separated from the reactants by gel filtration.

When insulin is [^{125}I]-labelled, the [^{125}I] isotope is incorporated into tyrosine residues according to their surface availability. Tyrosines exist at four positions: A14, A19, B16 and B26 (Figure 1.2). The resultant labelled insulin preparation is highly heterogeneous, consisting mainly of monoiodinated and diiodinated isomers, insulin polymers (resulting from decay of the latter) and unlabelled native insulin (Gammeltoft, 1984). Early endocrinological and immunochemical studies had shown that the biological activity of such iodinated mixtures in the unfractionated state is low due to oxidation damage and multiple iodine substitution (Izzo *et al*, 1964; Brunfeldt *et al*, 1968; Garrat *et al*, 1972). Thus, even in the preliminary, developmental stages of radioligand binding work, a number of laboratories attempted to improve the labelling technique by the introduction of milder iodination agents such as lactoperoxidase for the suppression of oxidation damage (Thorell and Johansson, 1971) and the reduction of iodination to minimal levels so that monoiodoinsulin was produced (Freychet *et al*, 1971); furthermore, over the next few years, manipulation of reaction conditions such as pH, mole ratio of iodine to insulin, and insulin concentration to achieve similar ends became commonplace. Analytical studies of protocols such as these, reviewed and summarised by Gammeltoft (1984), have shown that under mild conditions the majority of iodine becomes incorporated into the A chain, probably because the tyrosines therein are more exposed than those in the B chain when insulin is dimerised. They have shown also that of the two possible A chain-substituted monoiodoinsulin species, the A14 derivative is fully biologically active while the A19 derivative has attenuated activity. These observations are fully compatible with structure-function investigations which suggest

that Tyr^{A19} but not Tyr^{A14} is part of the receptor-binding region of the insulin molecule (Pullen *et al*, 1976). In recent years, therefore, purified preparations of monoiodo[^{125}I-Tyr^{A14}]insulin have been used routinely in insulin-receptor binding assays. Separation of one labelled insulin from another may be accomplished by gel filtration (Keefer *et al*, 1981), high performance liquid chromatography (HPLC) (Seidah and Chretien, 1983) or gel electrophoresis (Linde *et al*, 1983); very commonly, however, the approriately labelled insulin, already purified, is purchased from a commercial source. For the binding studies reported here, a commercially prepared human [^{125}I-Tyr^{A14}]insulin synthesised by lactoperoxidase iodination and purified by HPLC was used.

In addition to a radiolabelled ligand, any assay of ligand-receptor binding requires a population of intact cells, purified membranes or solubilised receptors. In the present work, cultured HepG2 cells were used as the source of insulin receptors and assays were conducted on whole cell populations.

HepG2 is a human liver cell line isolated by Knowles and coworkers (Aden *et al*, 1979) from a teenage Caucasian male from Argentina with primary hepatoblastoma and hepatocellular carcinoma. It was initially derived from minces of hepatoma biopsy tissue placed on feeder layers of the mouse cell line STO (Martin and Evans, 1975), irradiated to the point of non-division. Hepatoma sublines were selected after several months by feeder-independent *in vitro* growth in minimal essential medium (MEM) supplemented with foetal calf serum. After serial passage in MEM more than 50 times, HepG2 cells were reported to have recurrent chromosome abnormalities (between 50 and 56 chromosomes; mean = 55) with a distictive rearrangement of chromosome 1; in addition they were shown to retain the morphology and many of the biosynthetic capabilities of normal liver parenchyma (Knowles *et al*, 1980). The expression of differentiated functions is one of the most interesting and useful features of the HepG2 line. An extensive catalogue of

liver-specific phenotypic characteristics now known to be expressed in HepG2 cells has been tabulated by Darlington (1987). Integrated viral DNA is absent from the line, and Hepatitis B surface antigen is not synthesised. As a permanent cell line having many of the attributes of normal liver cells, it is not surprising that HepG2 carries a numerically significant population of insulin receptors, and has been chosen for a number of studies on these grounds (Ciechanover *et al*, 1983; Jacobs and Cuatrecasas, 1986). The HepG2 line was thus a strong contender for the insulin binding studies presented in this chapter. Use of a permanent cultured human cell line was viewed as preferable to the use of isolated hepatocytes (Gammeltoft *et al*, 1978) or adipocytes (Kono, 1975), which can vary greatly from one preparation to another, and must for convenience be derived from rodent models more distantly related to the insulin source species.

All receptor binding assays require, in addition to radioactively labelled ligand and a suitable source of receptors, a technique for the separation of bound and free ligand to be applied after each binding incubation. Separation techniques suitable for use with whole cells are limited; those applicable to cells in suspension include filtration (Levitzski, 1985) and centrifugation (Gammeltoft *et al*, 1978; Assoian and Tager, 1981). In the case of attached cells in culture, separation simply necessitates removal of the assay medium followed by washing and solubilisation of the cell fraction (Ciechanover *et al*, 1983; Haring *et al*, 1984). Since HepG2 cells grow adherent to the substratum, this last method was employed in the HepG2 binding assays reported here.

Insulin binds to its receptor in a reversible bimolecular reaction with high rate constants, a proportion of the bound ligand dissociating intact from the receptor at the cell surface. The insulin-receptor complex is endocytosed in an energy-requiring step. This, or one of the subsequent intracellular steps, is rate limiting, giving rate constants for

overall insulin uptake about one order of magnitude lower than those for the ligand-receptor binding reaction on its own (Sonne, 1988). Receptor binding assays designed to investigate the nature and kinetics of the bimolecular binding reaction in whole cells must, therefore, be conducted under conditions which inhibit endocytosis. The most convenient controlling parameter is temperature. Adsorptive endocytosis of insulin is known to be inhibited by temperatures below 16°C (Marshall, 1985). Accordingly, in the present work, prior to all binding procedures cells were washed to remove unbound extracellular insulin, incubated at 37°C for 1h in serum-free medium in order to clear the receptors of bound insulin, then cooled. Binding incubations were then set up in serum-free medium at 10°C.

The joint principles underlying any assay of hormone-receptor binding are those of receptor saturability and receptor specificity (Roth, 1973; Cuatrecasas *et al*, 1975). In practice this means that the receptor binding of physiological levels of [^{125}I]-labelled insulin (for example) reaches a point of equilibrium with both time and [^{125}I]insulin concentration, and can be inhibited in the presence of an excess of unlabelled insulin which has a high affinity for the receptor, but not by similar concentrations of other peptide hormones or other non-hormone proteins which have low affinities for the receptor. It is clear, therefore, that competitive inhibition provides a measure of specific receptor binding potential in the competitor species. In the work described in this chapter, preliminary experiments were carried out to establish the normal equilibrium characteristics of the binding of [^{125}I-Tyr^{A14}] human insulin to HepG2 cells. Subsequently, the capacity of various unlabelled substances to compete with free [^{125}I-Tyr^{A14}] human insulin for receptor sites on HepG2 cells was investigated in order to (i) establish insulin-receptor binding specificity and (ii) measure the specific receptor binding capacities of insulin-albumin conjugates.

4.2 MATERIALS AND METHODS

4.2.1 *MATERIALS*

HepG2 cells were a gift from Professor Gerry Coetzee, Department of Medical Biochemistry, University of Cape Town. Cell culture media were obtained from GibCo, Grand Island, New York, U.S.A., and sterile plasticware from Sterilin Ltd., Teddington, Middlesex, England. Human [^{125}I-Tyr^{A14}]insulin (specific radioactivity 2000 Ci/mmol) was purchased from Amersham International, U.K. All other chemicals were of analytical grade.

4.2.2 *METHODS*

Growth and Maintenance of HepG2 Cells

HepG2 cells were grown at 37°C in closed gas-tight flasks containing Eagle's minimal essential medium supplemented with 10mM $NaHCO_3$ and buffered with 20mM 2-[4-(2-hydroxyethyl)-1-piperazinyl]-ethanesulphonic acid-NaOH (Na-HEPES), pH 7.3 (MEM) containing 10%v/v foetal calf serum (MEM+S). Cells at semi-confluence (4-7 days' growth) were trypsinised with a solution containing 0.25%w/v trypsin, 0.1%w/v EDTA, and the suspensions subcultured at 1/6 or 1/3 dilution.

Protein Determination

Protein was estimated quantitatively by the method of Lowry (1951) as mentioned in section 2.2.2.

Receptor Binding Studies

HepG2 cells were heavily seeded (1/3 splits) into 35mm diameter well plates and grown to semi-confluence. Cells were washed

twice with MEM (no serum) at 37°C (2ml per well) then incubated at 37°C for 1h in the presence of a further 2ml MEM, which was subsequently removed. Plates were placed on ice and MEM at 5°C added (1ml per well) together with [^{125}I]insulin and unlabelled insulin or insulin conjugate as indicated in the figure legends. Additives were thoroughly mixed into the binding medium and plates incubated at 10°C. Cells were subsequently washed three times with cold phosphate buffered saline (PBS) (4ml per well) and drained. After addition of water (1ml per well), cells were loosened into suspension with a rubber policeman, 200µl removed for protein determination, and the remainder fully dispersed by the addition of 1ml lysis buffer (0.5% SDS, 100mM NaCl, 40mM Tris-HCl, 20mM EDTA, pH 7.0) (Shih and Weinberg, 1982). The lysates were transferred to tubes for gamma counting.

4.3 RESULTS

Characteristics of the binding interaction between human [^{125}I-Tyr^{A14}]insulin and HepG2 cells are shown in Figure 4.1. Figure 4.1A illustrates the time course of binding at 10°C. Binding took place rapidly for the first 2h, after which it slowed down, reaching a maximum at approximately 16h. Thus, for all subsequent binding experiments, incubation was for an 18h period at 10°C. The effect of insulin concentration on the binding response is presented in Figure 4.1B. The amount of insulin bound increased with insulin input, and at an insulin concentration of 0.8ng/ml, saturation of the receptor population in the system had not yet been reached.

Results of experiments designed to measure binding competition between unlabelled substances and human [^{125}I-Tyr^{A14}]insulin are shown in Figure 4.2. They demonstrate that porcine insulin and insulin-albumin conjugates derived from porcine insulin competed effectively with human insulin for receptor sites on HepG2 cells. Binding resulting from the addition of 0.4ng human

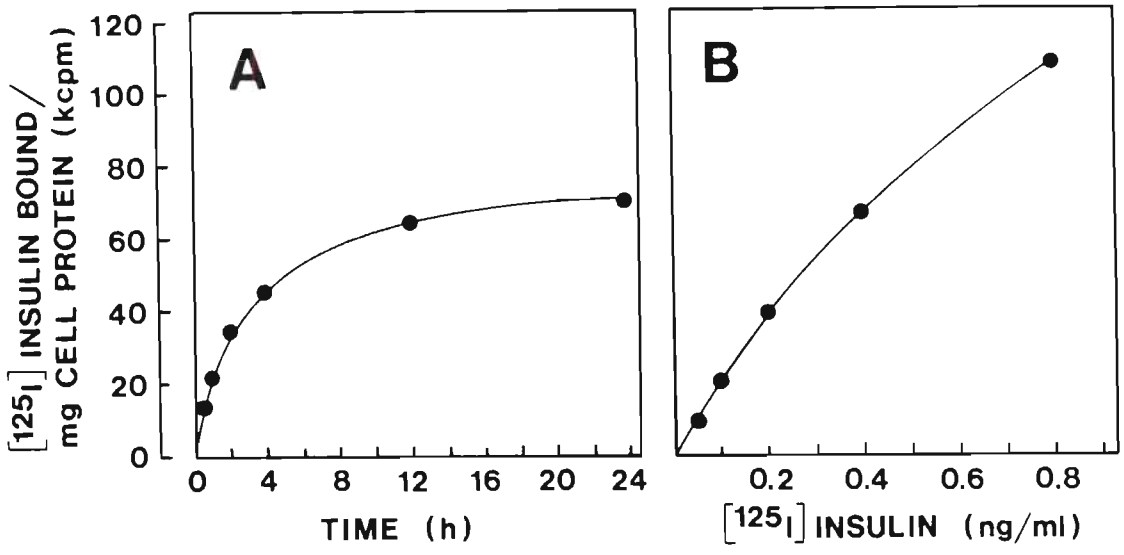


Figure 4.1 Binding of human insulin to HepG2 insulin receptors

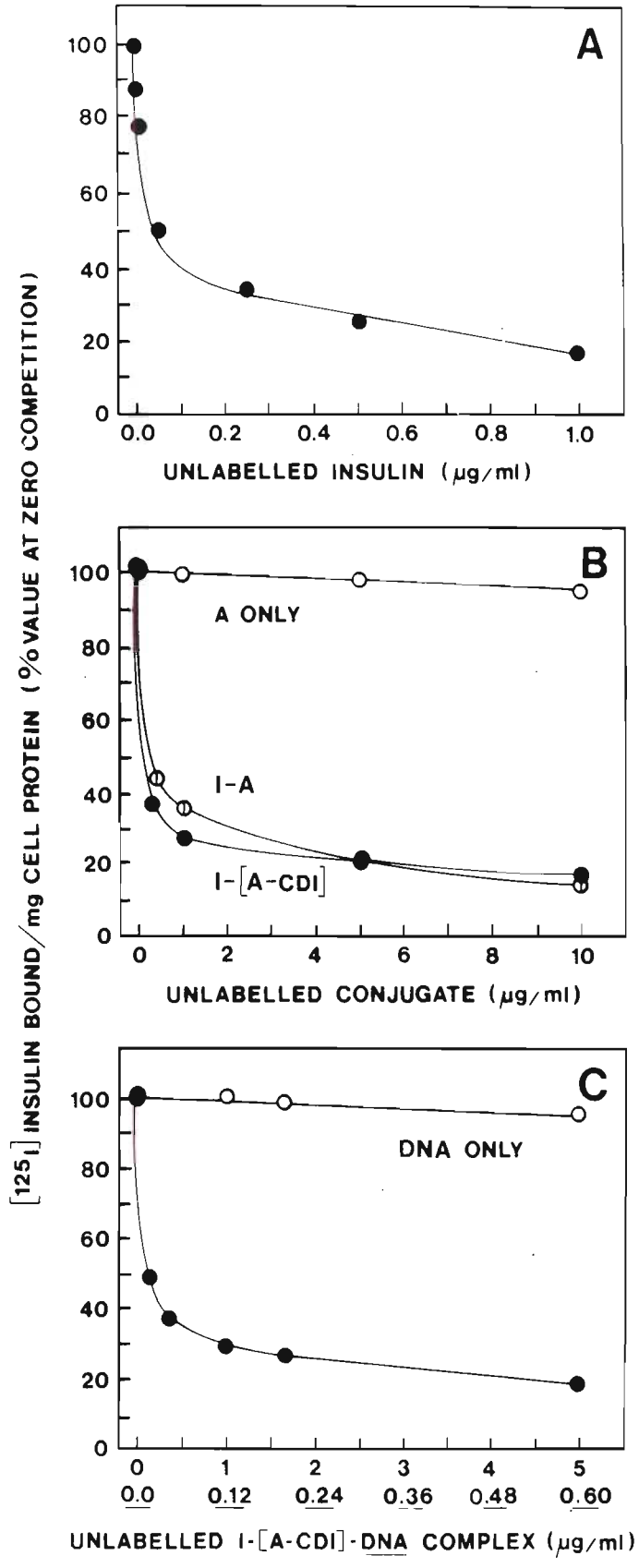
Semi-confluent monolayers of HepG2 cells were grown and prepared for binding as described in Materials and Methods.

A. Human[^{125}I-Tyr^{A14}]insulin (2µl, 0.4ng, 2.36 X 10^5 cpm) was added to each well and the incubation period at 10°C varied.

B. Variable amounts of human [^{125}I-Tyr^{A14}]insulin were added to the wells and incubation at 10°C allowed to continue for 18 hours.

After incubation each well was processed as described in Materials and Methods.

Figure overleaf/...


A
100
80
60
40
20
0
0.0
0.2
0.4
0.6
0.8
1.0
UNLABELLED INSULIN (μg/ml)
B
A ONLY
I-A
I-[A-CDI]
0
2
4
6
8
10
UNLABELLED CONJUGATE (μg/ml)
C
DNA ONLY
0
1
2
3
4
5
0.0
0.12
0.24
0.36
0.48
0.60
UNLABELLED I-[A-CDI]-DNA COMPLEX (μg/ml)
[125I] INSULIN BOUND/mg CELL PROTEIN (% VALUE AT ZERO COMPETITION)

[^{125}I-Tyr^{A14}]insulin alone was reduced to 50% in the presence of 50ng unlabelled porcine insulin (Figure 4.2A), 300ng unlabelled I-A and 200ng unlabelled I-[A-CDI] (Figure 4.2B). By contrast, albumin unconjugated to insulin did not compete for binding (Figure 4.2B). Furthermore, when I-[A-CDI] was prebound to DNA (pBR322), its receptor binding facility was not affected; the DNA itself had no affinity for the binding site (Figure 4.2C).

4.4 SUMMARY AND DISCUSSION

The binding of human insulin to HepG2 cells has been investigated in preliminary experiments (Figure 4.1), the results of which demonstrate the typical ligand-receptor binding characteristics of time-dependency (Figure 4.1A) and dependence upon ligand concentration (Figure 4.1B). Competitive displacement of human insulin by porcine insulin (Figure 4.2A) but not by albumin (Figure 4.2B) indicates that the binding reaction shown by insulin occurs specifically in relation to the insulin receptor and not at non-specific membrane binding sites. In the light of the non-competitive behaviour of albumin, the capacity of insulin-albumin conjugates to compete with human insulin (Figure 4.2B) demonstrates that the insulin constituents of the conjugates are capable of specific receptor recognition, and that the presence of *N*-acylurea substituents on the albumin component of I-[A-CDI] does not adversely affect conjugate- receptor binding affinity. Moreover, when I-[A-CDI] is allowed to form a complex with DNA prior to the cell binding incubation its receptor recognition facility is not impaired (Figure 4.2C).

In an investigation into the internalisation and recycling of the asialoglycoprotein receptor in relation to the activities of the transferrin and insulin receptors in HepG2 cells, Lodish and his coworkers (Ciechanover *et al*, 1983) have reported insulin-receptor binding characteristics which are fully

compatible with the results shown in Figure 4.1. The authors found that at 4°C binding was time-dependent, saturation of most surface sites being reached after 2h, and that binding was dependent on [^{125}I]insulin concentration, being inhibited by an excess of unlabelled insulin but not bovine serum albumin. Their binding curves are of interest in that, using 35mm diameter culture dishes and cells grown to semi-confluence as in the present work, receptor saturation appears not to have been fully reached even at 200nM [^{125}I]insulin (1140ng/ml). The apparent contrast between those results and the results in Figure 4.1 may possibly be explained by the fact that Ciechanover and colleagues iodinated their insulin "in house" and did not subject the reaction mixture to extensive purification.

For the receptor binding studies using the HepG2 cell line, Lodish's group used porcine insulin, while in the present work a combination of human and porcine insulins has been used. The structures of human and porcine insulins are almost identical; the only variation exists at position B30 which is occupied by threonine in the human, and alanine in the pig insulin (Smith, 1972). As a result of the sequence homology, the two insulins show identical binding properties in relation to receptors from a particular source (Gammeltoft, 1984).

The ability of insulin-albumin conjugates to bind insulin receptors as shown by competitive displacement of [^{125}I]insulin (Figure 4.2B) is rather marked. Only 300ng unlabelled I-A or 200ng unlabelled I-[A-CDI] is required to bring about the same degree of binding inhibition (50%) incurred by 50ng unlabelled insulin. While the molecular weight range of the conjugates is not known with any precision, it is certain that it is in the region of several hundred thousand (section 2.4), many times the molecular weight of monomeric insulin (5700), the reactive form of insulin under assay conditions. It would appear, therefore, that the affinity of conjugated insulin-albumin for the insulin receptor is greater (K_d value lower) than that of insulin itself

by an order of magnitude.

A similar phenomenon has been noted by several laboratories concerned with measuring the biological effects of macromolecular insulin derivatives. Cuatrecasas (1969) showed that when insulin is covalently attached to the insoluble polymer Sepharose, its action on isolated fat cells is amplified, and Suzuki and colleagues (1972) reported that a soluble purified form of insulin-dextran complex is more active than insulin itself both *in vivo* and *in vitro*. In both those investigations, evidence based on polymer size was provided to support the idea that the insulin conjugates remained exterior to the cell, and it was argued that the binding event at the receptor site alone, disassociated from any process of insulin translocation across the membrane, provided the stimulus for the hormone-induced biochemical responses observed. In a later study, Oka and Topper (1974) coined the epithet "super-active" for a soluble form of insulin, released from insulin-Sepharose in the presence of bovine serum albumin, which stimulates various insulin-induced biochemical processes in murine mammary tissue and cells five times more effectively than native insulin. The structure of the "super-active" insulin was revealed in a later study by the same group (Wilchek *et al*, 1975). The insulin-Sepharose had been synthesised by cyanogen bromide activation of the Sepharose followed by coupling via an amino group on the insulin molecule. On the basis of work with model compounds, such conjugates were shown to be *O*-Sepharose-*N*-substituted isoureas capable of further nucleophilic attack by amino groups to give soluble *N*-*N'*-disubstituted guanidines unattached to the Sepharose. The "super-active" insulin was postulated to be a guanidine derivative of this kind in which insulin and albumin are the substituents: a small conjugate relative to the insulin-Sepharose, quite able to enter cells through the membrane. The authors point out that their elucidation of the chemistry of insulin derivatives of this nature might necessitate reexamination of the interpretation previously

placed on results obtained with peptide hormone-Sepharose conjugates. The conjugates used by Cuatrecasas (1969), for example, were also synthesised by cyanogen bromide activation of Sepharose and therefore could have undergone the reactions described by Wilchek's group in the presence of cell buffering media containing BSA or other proteins.

Wilchek and coauthors state that the role of the disubstituted guanidine residue in conferring super-activity on a hormone such as insulin is unknown. The apparent functional similarity between their guanidine-linked conjugates and the glutaraldehyde-linked conjugates used in the present work suggests that the linking structures themselves might be of minor significance. It is possible, rather, that some aspect of insulin structure altered by the conjugation process is responsible for the enhanced activity.

Certain insulins modified at particular amino acid residues have been shown to display increased binding affinities, and the results of studies conducted on these compounds allow considerable insight into the structure-function relationships involved in binding affinity variations. Compared with unmodified insulin, monoiodo[Tyr^{B26}]insulin demonstrates 200% and [D-Phe^{B24}]insulin 180% increased receptor affinity. In the case of monoiodo[Tyr^{B26}]insulin the raised affinity (lowered K_d) is explained by an increase in the rate of association (Sonne *et al*, 1983) while in the case of [D-Phe^{B24}]insulin a decrease in the rate of dissociation occurs (Kobayashi *et al*, 1982). In a manner similar to the latter, chicken insulin shows a decrease in the rate of dissociation from human insulin receptors and is thus seen to display enhanced binding (Simon *et al*, 1977). Altered binding affinities such as these may have fairly simple causes. In the cases of the modified Phe^{B24} and Tyr^{B26} residues, which are part of the receptor binding region (Figure 1.4), the minor changes in structure clearly have direct implications in the binding interaction, possibly in the hydrophobic behaviour of the cluster of B chain residues involved. In chicken

insulin, a histidine is present at the A8 position, just outside the receptor binding region, which in bovine and porcine insulins is occupied by alanine and in human insulin by threonine (Smith, 1972). Histidine, having a larger side group than alanine or threonine, may form stabilising interactions with amino acid residues in the receptor or induce minor rearrangements of residues in the receptor binding region of the ligand (Pullen *et al*, 1976), thereby inducing an alteration in binding affinity. Like chicken insulin, a number of fish insulins are characterised by His^{A8} (Gammeltoft, 1984). In discussing the biological activity of Atlantic hagfish insulin in relation to structure, Cutfield and coworkers (1979) make the notable point that a change in K_d by a factor of 2 or 3 corresponds to only a small difference in standard free energy (ΔG°): of the order of 1 - 2 kcal/mol. A shift of this magnitude may be due to a difference of just a few hydrogen bonds in the insulin-receptor binding reaction, and could result from minor conformational changes.

It was reasoned previously (section 2.4) that glutaraldehyde linkage of insulin to another protein probably occurs through the insulin Lys^{B29} residue, which is not within the deduced receptor binding site but is quite close to it (Figure 1.4), and it was argued that such a linkage would be unlikely to interfere with receptor binding. Results shown in Figure 4.2 support the concept of linkage outside the binding site, but do suggest that indirect modifications might have been induced at the binding site itself. Such rearrangements, if they do exist, might be analogous to those thought to be induced by the His^{A8} residues of chicken and fish insulins, since the B29 and A8 amino acids occupy positions similarly close to the periphery of the binding site (Figure 1.4).

The work presented in this chapter is largely concerned with establishing the qualitative nature of the binding response between insulin-albumin conjugates and HepG2 cells, not with precise measurement of binding coefficients. As Sonne (1988)

has emphasised, kinetic values obtained from experiments performed at low temperature to inhibit endocytosis are valid only for that temperature: binding of insulin to IM-9 lymphocytes in which receptor-linked endocytosis is absent has been shown to be markedly temperature dependent (Waelbroeck *et al*, 1979), and comparable temperature effects have been demonstrated in fat cells and liver cells (Gammeltoft, 1984). Binding affinity is higher at lower temperatures. It is clear, therefore, that conjugate-receptor binding as observed at 10°C (Figure 4.2) is not going to be quantitatively the same under transfection conditions at 37°C (Chapter 5); there is likely to be a shift towards lower affinity (higher K_d values). Nevertheless, the assumption is made here that the binding responses of insulin and its conjugates to HepG2 cell receptors bear similar relationships to one another regardless of temperature, and that the general effects illustrated in Figure 4.2 hold true at 37°C.

In the experiment designed to investigate binding of the I-[A-CDI]-DNA complex to the receptor (results shown in Figure 4.2C) a situation existed which, it was clear, would be encountered again later in transfection work (Chapter 5): that of a disparity between optimal conditions for conjugate-DNA binding (simple salt buffer; ionic strength, 50mM or less; osmolality <100 mmol/kg) and suitably protective conditions for conjugate-cell receptor binding (complex cell culture medium; osmolality, 298 mmol/kg). The experimental procedure is described in the legend to Figure 4.2C. The I-[A-CDI] conjugate was prebound to pBR322 DNA under precisely controlled buffer conditions designed to give a final concentration of 50mM NaCl and a final pH of 6.5. The complex was then added to MEM as part of the regular receptor binding protocol. On account of its composition, MEM was thought likely to cause an equilibrium shift towards dissociation of DNA from the I-[A-CDI] in a manner similar to Na^+ as illustrated in Figure 3.5, thus reducing the proportion of DNA bound to *N*-acylurea protein at the time of receptor binding. Nitrocellulose filter binding assays

indicated that this was indeed the case, the amount of DNA bound to I-[A-CDI] in buffer being somewhat reduced on MEM addition in the appropriate proportions (results not shown). However, on the basis that the transfection procedure involves the use of MEM medium and closely parallels the receptor binding protocol, the results presented in Figure 4.2C do validly demonstrate that DNA does not interfere with I-[A-CDI]-receptor binding under the proposed transfection conditions.

CHAPTER FIVE

INSULIN-[*N*-ACYLUREA ALBUMIN] MEDIATED GENE TRANSFER TO HepG2 CELLS AND SUBSEQUENT EXPRESSION

5.1 INTRODUCTION

The reporter gene chosen for testing the ability of insulin-[*N*-acylurea albumin] to mediate transfection in HepG2 cells was the stable dominant marker *neo* (Colbere-Garapin *et al*, 1981), which can be detected by means of the selective growth of recipient cells in medium containing G418 at a concentration lethal to normal cells (section 1.2.1).

The compound G418 is one of a large family of aminoglycoside-aminocyclitol antibiotics characterised by the cyclitol 2-deoxystreptamine. This group includes the neomycins, in which the cyclitol is 4,5-disubstituted, as well as the kanamycins, gentamicins and tobramycin, in which the cyclitol is 4,6-disubstituted. G418 belongs to the latter class, and is structurally most closely related to the gentamicins, a large, extremely important group of antibiotic compounds produced by strains of the actinomycete genus *Micromonospora* (Hooper, 1982). The commercial material termed gentamicin consists primarily of gentamicins C_1, C_2 and C_{1a}, products of *Micromonospora purpurea*, first reported by Weinstein *et al* (1964), the structures of which were elucidated later by workers of the Schering Corporation (Cooper *et al*, 1971). Several naturally occurring minor components of the gentamicin complex obtained by the fermentation of *M. purpurea* and other closely related cultures have been isolated and characterised, including the gentamicin A group (Maehr and Schaffner, 1967), gentamicins B and B_1, gentamicin X_2, and the antibiotic G418 (Wagman *et al*, 1972;

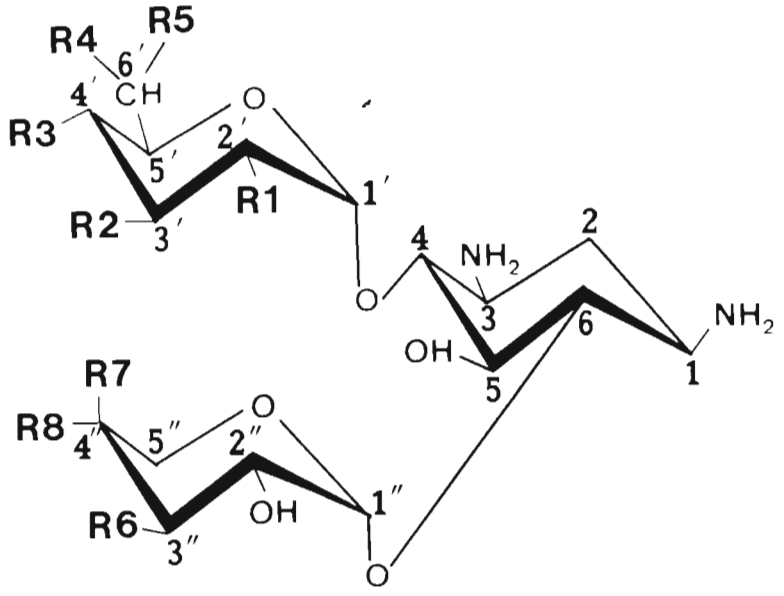
Berdy *et al*, 1977). Structural analysis of G418 was accomplished by the Schering Corporation research group (Daniels *et al*, 1973). The structural characteristics of G418 in relation to other important gentamicin-like compounds are provided in Figure 5.1.

All the aminoglycoside antibiotics are active against a range of prokaryotic organisms. They interfere with ribosome function, and are usually bactericidal, although a few are merely bacteriostatic (Tanaka, 1982). In addition to their action on bacterial protein synthesis, a small number affect certain eukaryotic organisms in a similar manner by interacting with mitochondrial ribosomes (Kurtz, 1974), which have architectural and functional features in common with the prokaryote ribosome (Margulis, 1974). The aminoglycoside antibiotics which promote eukaryotic translation errors in this way include paromomycin, lividomycin B, hygromycin B, as well as the 2-deoxystreptamine antibiotics kanamycins B and C (Wilhelm *et al*, 1978a,b; Singh *et al*, 1979) and G418 (Jimenez and Davies, 1980). The antibiotic G418 has a particularly broad spectrum of activity in both prokaryotes and eukaryotes: it is an extremely potent inhibitor of protein synthesis in cell-free extracts of bacterial, fungal, algal, plant and animal cells (S. Perzynski, unpublished observations, quoted in Jimenez and Davies, 1980). Perzynski's results are of especial interest in that they infer the ability of G418 to act on cytoplasmic as well as organellar ribosomes.

The most common form of resistance to aminoglycoside antibiotics is provided by plasmid-encoded modifying enzymes which may be classified according to mode of action (*N*-acetylation, *O*-nucleotidylation or *O*-phosphorylation) and site of action on the antibiotic. In this manner twelve major classes of enzyme have been recognised. They have been tabulated by Davies and Smith (1978) and Umezawa and Kondo (1982), the latter authors emphasising the further class subdivisions which have been made in some cases to indicate diverse substrate specificity. A summary of the basic classification scheme is provided in Table

5.1, and the example of class subdivision most relevant to the present work shown in Table 5.2. Table 5.1 illustrates that deoxystreptamine aminoglycosides are variously subject to modification by all three major classes of *N*-acetyltransferases (AAC), two out of four classes of *O*-nucleotidyltransferases (AAD), and two out of five classes of *O*-phosphotransferases (APH). Enzymes of the APH class, the APH(3′) group, are the ones encoded by *neo* sequences. The site of phosphorylation of this class of enzyme is the 3′-hydroxyl of the 4,-substituent amino-hexo- or hepto- pyranose (Table 5.1; Figure 5.1). The APH(3′) enzymes are therefore capable of transferring phosphate groups to kanamycins, neomycins, and other 2-deoxystreptamine antibiotics having hydroxyl groups at the 3′ position and were originally called neomycin-kanamycin phosphotransferases (Umezawa, 1974), hence the designation *neo* for the corresponding genes. These same enzymes are now termed aminoglycoside 3′-phosphotransferases (Umezawa and Kondo, 1982). As shown in Figure 5.1, the antibiotic G418 is characterised by a 3′-hydroxyl group and, as shown in Table 5.2, is susceptible to phosphorylation by APH(3′) enzymes.

Aminoglycoside 3′-phosphotransferase enzymes are of three types, reflecting minor differences in substrate specificity: APH(3′)I, APH(3′)II and APH(3′)III (Table 5.2). Types I and II are both capable of phosphorylating the antibiotic G418. APH(3′)I is found in many different bacterial genera, representing both Gram positive and Gram negative types (Davies and Smith, 1978). It appears to exist in a number of forms which demonstrate variations in molecular weight, pH optimum and K_i according to source (Matsuhashi *et al*, 1975). In some organisms the enzyme is encoded on a transposable element rather than a plasmid. It was APH(3′)I from the transposon Tn601(903) (Davies *et al*, 1977) which was used to demonstrate the potential of the *neo* selection system in eukaryotes for the first time: sphaeroplasts of *Saccharomyces cerevisiae* were transformed with a plasmid carrying the transposon and screened for the expression of the APH(3′)I resistance determinant on G418 gradient agar plates



		R1	R2	R3	R4	R5	R6	R7	R8
"Gentamicin"	Gentamicin C_1	NH_2	H	H	CH_3	$NHCH_3$	$NHCH_3$	OH	CH_3
	C_2	NH_2	H	H	CH_3	NH_2	$NHCH_3$	OH	CH_3
	C_{1a}	NH_2	H	H	H	NH_2	$NHCH_3$	OH	CH_3
Gentamicin-related antibiotics	Gentamicin A	NH_2	OH	OH	H	OH	$NHCH_3$	H	OH
	A_1	NH_2	OH	OH	H	OH	$NHCH_3$	OH	H
	A_2	NH_2	OH	OH	H	OH	OH	H	OH
	A_3	OH	OH	OH	H	NH_2	$NHCH_3$	OH	H
	Gentamicin B	OH	OH	OH	H	NH_2	$NHCH_3$	OH	CH_3
	B_1	OH	OH	OH	CH_3	NH_2	$NHCH_3$	OH	CH_3
	Gentamicin X_2	NH_2	OH	OH	H	OH	$NHCH_3$	OH	CH_3
	Antibiotic G418	NH_2	OH	OH	CH_3	OH	$NHCH_3$	OH	CH_3

□ = variant moieties within related groups

Figure 5.1 The structure of the antibiotic G418 in relation to the gentamicin family of 4,6-disubstituted 2-deoxystreptamine aminoglycoside antibiotics (after Okuda and Ito, 1982, and Hooper, 1982).
Trade name of G418: Geneticin
Full chemical name of G418: *O*-2-Amino-2,7-dideoxy-D-glycero-α-D-glucoheptopyranosyl [1-4][-*O*-3-deoxy-4C-methyl-3-[methylamino]-ß-L-arabinopyranosyl]-D-streptamine.

Table 5.1 Classes of aminoglycoside-modifying enzymes (after Davies and Smith, 1978).

AS I and II = amino sugars as shown in Figure 5.1
AG = aminoglycoside

TRANSFERASE ENZYME	ABBREVIATION	SITE OF MODIFICATION
3-*N*-acetyl-	AAC(3)	3-NH_2 on deoxystreptamine
2′-*N*-acetyl-	AAC(2′)	2′-NH_2 on AS I of deoxystreptamine AGs
6′-*N*-acetyl-	AAC(6′)	6′-NH_2 on AS I of deoxystreptamine AGs
6-*O*-nucleotidyl-	AAD(6)	6-OH on streptidine ring of streptomycin
4′-*O*-nucleotidyl-	AAD(4′)	4′-OH on AS I of deoxystreptamine AGs
2″-*O*-nucleotidyl-	AAD(2″)	2″-OH on AS II of deoxystreptamine AGs
3″(9)-*O*-nucleotidyl-	AAD(3″)(9)	3″-OH on AS II of streptomycin or 9-OH on actinamine of spectinomycin
6-*O*-phospho-	APH(6)	6-OH on streptidine ring of streptomycin
3′-*O*-phospho-	APH(3′)	3′-OH on AS I of deoxystreptamine AGs
2″-*O*-phospho-	APH(2″)	2″-OH on AS II of deoxystreptamine AGs
3″-*O*-phospho-	APH(3″)	3″-OH on AS II of streptomycin
5″-*O*-phospho-	APH(5″)	5″-OH on ribose of ribostamycin

Table 5.2 Substrate specificities of aminoglycoside 3′-phosphotransferases (After Umezawa and Kondo, 1982; additional information from Colbere-Garapin *et al*, 1981).

ENZYME	KNOWN SUBSTRATES
APH(3′)I	KM, KM-B, NM, PM, RM, **G418**, LV(5"-OH).
APH(3′)II	KM, KM-B, NM, PM, RM, **G418**, BT.
APH(3′)III	KM, KM-B, RM, BT, LV(5"-OH).

Abbreviations: KM, kanamycin; NM, neomycin; PM, paromomycin; RM, ribostamycin; LV, lividomycin; BT, butirosin.

Both APH(3′)I and II are able to phosphorylate the 3′-OH of kanamycin, kanamycin B, neomycin, paromomycin, ribostamycin and G418. The contrast in substrate specificity lies in the fact that APH(3′)I can phosphorylate the 5"-OH group of the lividomycins but not the 3′-OH of the butirosins (which have a 4-amino-2-hydroxybutyryl group as a substituent on the 1-amino), while APH(3′)II shows the reverse characteristic. No information could be found on the susceptibility of gentamicins A series, B series and X_2 to APH(3′) enzymes.

(Jimenez and Davies, 1980). In contrast with APH(3')I, APH(3')II occurs less widely, but has been detected in at least eight bacterial genera and there is evidence to suggest that, regardless of source, only one form of the enzyme exists (Matsuhashi *et al*, 1975). It has a molecular weight of 27,000 (Matsuhashi *et al*, 1976), a K_m for ATP of 24 $\pm$ 4μM and a K_m for neomycin B of 3.9μM (Davies and Smith, 1978). Jorgensen (1979) discovered that a *neo* sequence coding for APH(3')II is carried on the transposable element Tn5, and it was the gene from this source, in a fragment defined by Rothstein *et al* (1980), which was used by Garapin's group (Colbere-Garapin *et al*, 1981) in the first assay of *neo* expression in mammalian cells. Garapin and coworkers set an important precedent in developing the APH(3')II mammalian expression assay, for it constitutes the basis of an enormous number of transfection studies undertaken in the past eight years, including those reported in this thesis.

For the gene transfer work described in this chapter, two expression vectors were employed, each incorporating *neo* from Tn5. The principles of construction of the two vectors, ptkNEO and pAL-8, are outlined in Figures 5.2 and 5.3 respectively. Each carries a pBR322 sequence which includes the origin of bacterial replication and the ampicillin resistance (*amp*^{R}) gene to facilitate selective growth of plasmid-containing cells during preparative plasmid production in a host bacterium (section 1.2.1). The essential difference between the two vectors lies in the nature of the eukaryotic regulatory sequences introduced into the constructs on either side of *neo*. The ptkNEO vector contains control regions taken from the herpes simplex virus Type 1 (HSV1) thymidine kinase (*tk*) gene (Wagner *et al*, 1981), while pAL-8 incorporates control sequences from simian virus 40 (SV40) (Reddy *et al*, 1978; Laimins *et al*, 1982). An additional difference between the two vectors, which makes a considerable distinction in terms of size, but little functional difference in the context of the experiments conducted here, is that the vector pAL-8 carries a further insert, a 6.6kb sequence from the plasmid pT24-C3 representing the oncogene *ras*.

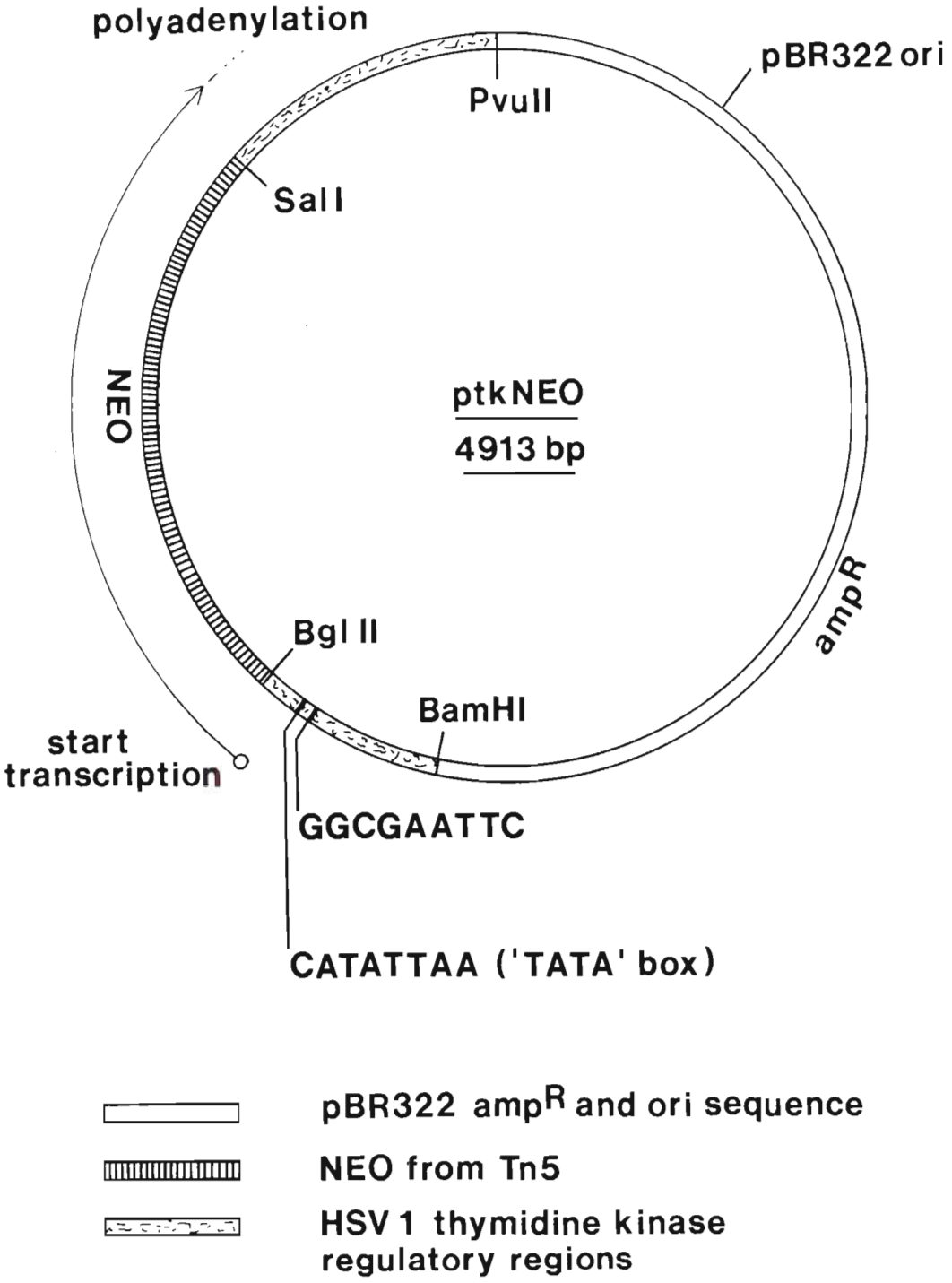


Figure 5.2 Construction of the expression vector ptkNEO
(Barbara Wold)
The full sequence is provided in Appendix I.

sequence		base position	size
PvuII*-BamHI	pBR322 *amp*^{R} and ori	1-2653	(2653bp)
BamHI-BglII	HSV1 *tk* 5′ proximal to *neo*	2654-3117	(464bp)
BglII-SalI	Tn5 *neo*	3118-4287	(1170bp)
SmaI*-PvuII	HSV1 *tk* 3′ proximal to *neo*	4288-4913	(626bp)

(* PvuII and SmaI sites destroyed)

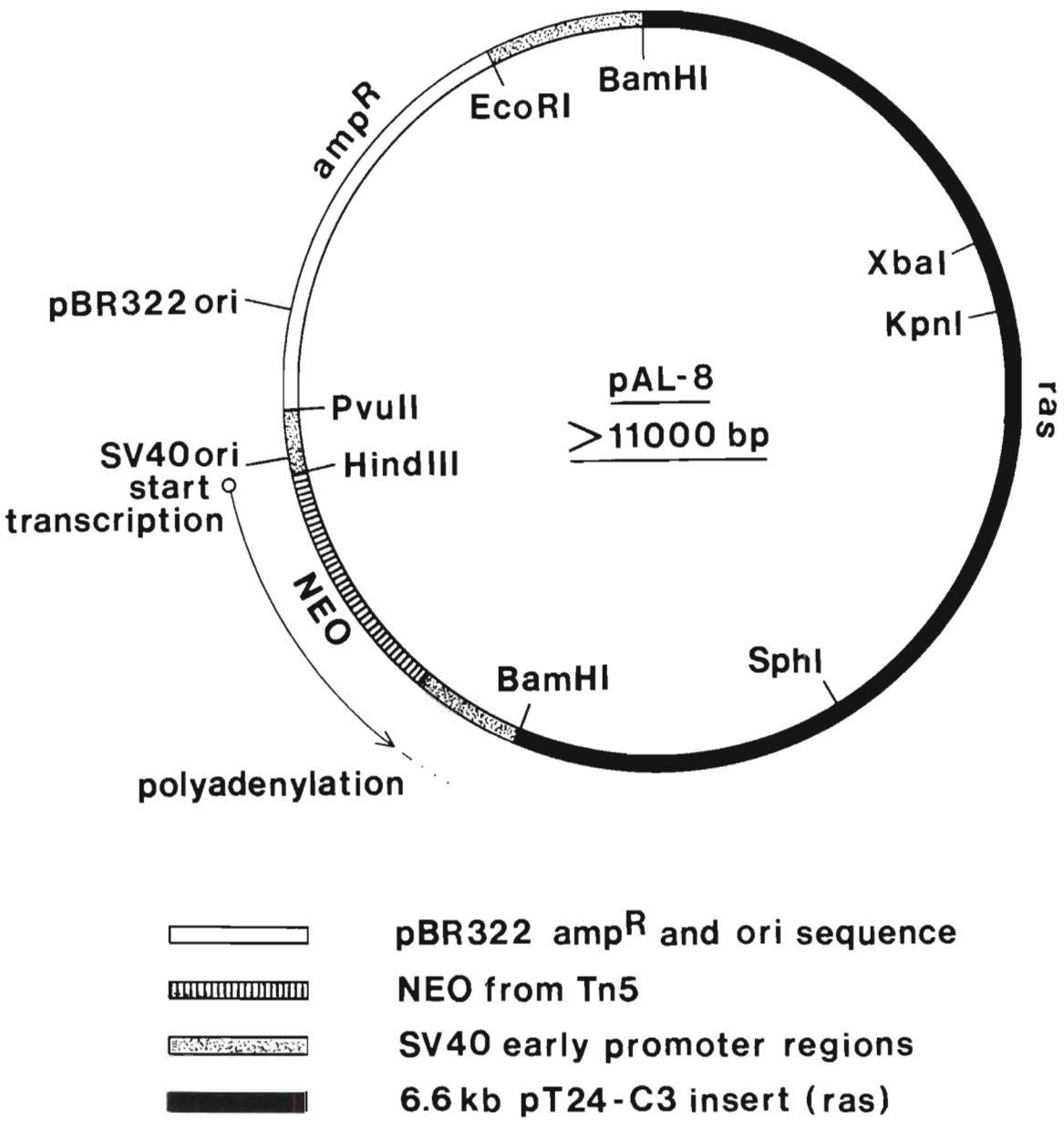


Figure 5.3 Construction of the expression vector pAL-8

The HSV1 *tk* gene regulatory regions employed in the ptkNEO vector (Figure 5.2) are utilised in much the same fashion as in the naturally occurring *tk* gene itself, as elucidated by Wagner and coworkers (1981) in a major analysis of the *tk* nucleotide sequence. In a position 5′ proximal to *neo* is inserted a 463bp BamHI/BglII fragment from upstream of the *tk* open reading frame (ORF) which includes a Goldberg-Hogness-like "TATA" box (CATATTAA), the first base pair of which determines the beginning of mRNA production 31 nucleotides further along the sequence. In addition, 44 nucleotides prior to the "TATA" box (85 nucleotides prior to the start of the mRNA), is a sequence (GGCGAATTC) of probable regulatory significance, similar to one noted by Benoist *et al* (1980) to be conserved in many eukaryotic genes. While the mRNA start position lies within the *tk* insert, the first ATG indicating the mRNA AUG codon for the initiation of translation is within the *neo* sequence itself. At the other end of *neo* (3′ proximal) is inserted a 620bp SalI/PvuII sequence from downstream of the *tk* ORF. This includes, close to the beginning, thus almost immediately following the *neo* sequence, the translation termination codon TGA (mRNA, UGA), followed by two identical short sequences (AATAAA), 48 and 61 nucleotides beyond the TGA respectively. These hexanucleotide regions constitute examples of the polyadenylation signal common to all polyadenylated mRNAs, occurring 14-30 nucleotides prior to the beginning of the poly(A) tail (Proudfoot and Brownlee, 1976).

The SV40 regulatory sequences employed in the construction of pAL-8 (Figure 5.3) appear to be much the same as those used in the classical pSV_2 vectors (Gorman, 1985) (section 1.2.1). Upstream of *neo* is the 323bp PvuII/HindIII fragment of the SV40 early region promoter which includes a 72bp repeated enhancer (Gruss *et al*, 1981; Laimins *et al*, 1982), a"TATA" box (Ghosh *et al*, 1981), the SV40 origin of replication and the site for the start of transcription (Laimins *et al*, 1982). Downstream of *neo* is a larger fragment of the SV40 genome which terminates at a BamHI site and thus very probably parallels the sequence utilised in a similar position in pSV_2 constructs. This

incorporates a splice donor-acceptor sequence in the form of the small t-antigen intron, followed by polyadenylation signals (Gorman, 1985). A more precise analysis of the region downstream from *neo* is not possible at the present time as a detailed map of the pAL-8 construct has not been made available from source. Information concerning the SV40 region inserted between the pBR322 sequence and *ras* is sparse for the same reason.

Prior to their use in transfection of HepG2 cells, it was necessary that both ptkNEO and pAL-8 be replicated in and harvested from a suitable host bacterium. The host chosen for this purpose was *Escherichia coli* strain HB101, a hybrid resulting from a cross between *E. coli* K-12 and *E. coli* B, commonly used as a plasmid recipient via transformation and regarded as a favourable host for large-scale plasmid production (Bolivar and Backman, 1979). Its genotype is F^-, leu^-, thi^-, *hsdS20* (r_B^-, m_B^-), *recA13*, *ara-14*, *proA2*, *lacY1*, *galK2*, *rpsL20* (Sm^r), *xyl-5*, *mtl-1*, *supE44*, λ^- (Bolivar and Backman, 1979; Maniatis, 1982).

The effective transfer of a ptkNEO or pAL-8 vector during the transfection procedure and subsequent stable expression of the *neo* gene in recipient HepG2 cells was seen to be observable in two ways. Firstly, and primarily, by long term phenotypic selection leading to the demonstration of cell survival during a period of G418 treatment shown to be inimical and ultimately lethal to the untransfected cell line. Secondly, genotypically, by the identification of the *neo* gene sequence in DNA extracted from stably G418-resistant cell populations derived from transfected cell cultures. Clearly, the second of these approaches requires a positive result in the first, and is therefore by its very nature corroborative. In the event of positive selection of G418 resistance, it was planned to follow this second approach by employing the Southern blotting technique (Southern, 1975), in which electrophoretically size-fractionated DNA is transferred from an agarose gel matrix

to a solid membrane support, and subsequently hybridised to a labelled DNA probe comprising the sequence of interest. The Southern blotting and hybridisation method, capable of the identification of single copy gene insertions, has been described as "one of the most powerful tools of molecular biology" (Wahl *et al*, 1987). Its extreme sensitivity and specificity at the DNA sequence level was regarded as particularly significant in the context of using the *neo*-G418 system to detect transfection: positive results would provide definitive, molecular proof of gene transfer events otherwise only phenotypically inferred.

5.2 MATERIALS AND METHODS

5.2.1 *MATERIALS*

The expression vector ptkNEO, constructed by Dr Barbara Wold of the Biology Division, California Institute of Technology, was supplied by Dr Shirley Taylor of the same institution. The expression vector pAL-8 was a gift from Professor Peter Jones, Comprehensive Cancer Center, University of Southern California School of Medicine. *Escherichia coli* strain HB101 was supplied by Professor Dave Woods, Microbial Genetics Research Unit, Department of Microbiology, University of Cape Town. Bacteriological culture media and ingredients were purchased from Difco Laboratories, Detroit, Michigan, U.S.A., cell culture media from GibCo, Grand Island, New York, U.S.A., and sterile plasticware from Sterilin Ltd., Teddington, Middlesex, England. Filter sterilisation units were from Millipore Corporation, Bedford, Mass, U.S.A. DNA polymerase I (EC 2.7.7.7) Klenow fragment, restriction endonucleases, λ DNA and plasmid pBR322 were obtained from Boehringer Mannheim GmbH, Penzberg, Western Germany. Lysozyme (EC 3.2.1.17) (Product Code No. L-6876), ethidium bromide (Product Code No. E-8751), ampicillin (Product Code No. A-9518), chloramphenicol (Product Code No. C-0378),

Geneticin (G418) (Product Code No. G-5013), polyvinylpyrrolidone (average mol.wt. 360000) (Product Code No. P-5288), dextran sulphate (average mol.wt. 500000) (Product Code No. D-8906) and ficoll (average mol.wt. 400000) (Product Code No. F-4375) were from the Sigma Chemical Co., St. Louis, Mo. U.S.A. Thymidine 5'-[α-^{32}P] triphosphate (specific radioactivity 3000 Ci/mmol), Hybond C-extra nitrocellulose blotting membrane and Hyper-film MP X-ray film were obtained from Amersham International, U.K. X-ray film developer (G150c) and rapid X-ray film fixer (G334c) were purchased from Agfa-Gevaert, Durban, R.S.A. Proteinase K (Code 24568) and caesium chloride (Code 2039) were obtained from Merck, Darmstadt, Federal Republic of Germany. Agarose (ultra pure grade and low melting temperature grade) was obtained from BioRad Laboratories, Richmond, Ca, U.S.A. Sephadex G-50 was from Pharmacia Fine Chemicals, Uppsala, Sweden. All other reagents were of analytical grade.

5.2.2 *METHODS*

Maintenance of HepG2 Cells

HepG2 cells were grown and maintained as described in section 4.2.2.

Generation of Expression Vectors

The expression vector plasmids ptkNEO and pAL-8 were each propagated in *Escherichia coli* strain HB101. Initial transformation was achieved using a simple calcium method (Harley *et al*, 1982) (section 1.1.2) and stocks of the respectively transformed host strain kept on ampicillin media as both anaerobic stab cultures and frozen glycerol-broth cultures. Large scale plasmid production followed conventional protocols (Maniatis *et al*, 1982) involving growth of the host to the late logarithmic phase in 650-1000ml volumes of aerated broth containing ampicillin (50µg/ml), followed by amplification of

the plasmid under the influence of chloramphenicol (170µg/ml) over a period of 16-18h, and subsequent harvesting of cells by refrigerated centrifugation at 4000Xg for 10min. Extraction of plasmids was by the boiling method of Holmes and Quigley (1981) or the cleared lysate method of Katz *et al* (1973), and purification was by centrifugation in caesium chloride-ethidium bromide (CsCl-EB) density gradients (Radloff *et al*, 1967). The CsCl-EB-DNA solutions (refractive index, η = 1.3860) were set up in open-topped nitrocellulose tubes, sealed with liquid paraffin, and centrifuged in a SW50 rotor at 40,000rpm for 40h at 20°C in a Beckman L5-65 ultracentrifuge. Plasmid bands were removed manually under UV illumination using drawn out pasteur pipettes inserted from the top of the tube. Removal of EB from the plasmid fractions was accomplished by several extractions with water-saturated redistilled *n*-butanol; final traces of butanol were removed by an ether extraction. The plasmid solutions were subsequently dialysed exhaustively at 4°C against TE buffer (10mM Tris-HCl, 1mM EDTA, pH 7.5). The final preparations were stored over chloroform (5%v/v) at -15°C. OD_{260}/OD_{280} values: 1.68 - 1.82. Yields: 900 - 950µg per 1000ml original culture.

Analysis and Quantification of DNA

Spectral analysis of DNA was accomplished using either a Beckman 24 or a Philips PU8700 series spectrophotometer in the UV range 220-320nm. Estimations of purity were made by comparing OD_{260}/OD_{280} ratios to the protein-free ideal (1.8), and concentrations calculated by applying the relationship: OD of 1.0 = 50µg double-stranded DNA/ml (Maniatis *et al*, 1982).

Agarose Gel Electrophoresis of Plasmid Vectors

Plasmids and restriction digests of plasmids were analysed by separation in 1.2%w/v agarose gels at neutral pH. A BioRad Mini Sub DNA cell system was used: gel dimensions 100 X 65 X 3mm. Gels were made and run in buffer containing 36mM Tris-HCl, 30mM

sodium phosphate and 10mM EDTA (final pH 7.5). Each well sample contained 0.4-0.6µg (2-7µl) DNA. Prior to gel application, samples consisting of linear restriction fragments were heated to 65°C for 10min, cooled on ice, then mixed with 1/3 the sample volume of loading buffer (50%w/v sucrose, 4M urea, 0.1%w/v bromophenol blue, 50mM EDTA, pH 7.5). Unrestricted plasmid samples were mixed with loading buffer without prior heat treatment. Electrophoresis was at 40-50V for 2-3h. Preparative electrophoresis for the separation and excision of restriction fragments for subsequent labelling as probes was performed in a similar manner on gels of low melting temperature agarose.

Visualisation of DNA Gels

Immediately following electrophoresis, DNA gels were stained in a dilute solution of ethidium bromide (0.5µg/ml; 15min), destained briefly in running water (±2 min), then viewed by transmitted UV light at 300nm wavelength. Ultraviolet illuminated gels were photographed using Kodak Tri-X Pan black and white film (ASA 400) in a single lens reflex camera fitted with a 55mm close-up lens and a red filter (exposure: f11, 20 - 35sec).

Labelling of neo Probe

A 923bp PstI fragment of ptkNEO representing the majority of the *neo* gene sequence and no extraneous sequences (Figure 5.5) was labelled with thymidine 5'-[α-^{32}P] triphosphate by the method of Feinberg and Vogelstein (1983) in which the denatured DNA is allowed to anneal in the presence of random hexanucleotide primers and polymerise under the influence of the DNA polymerase I (Klenow) enzyme. The desired plasmid sequence was excised from an EB-stained electrophoretic gel of low melting temperature agarose, water added (3µl/mg gel), and the sample boiled for 7 minutes immediately prior to use in the labelling reaction (Feinberg and Vogelstein, 1984). The reaction product was subjected to spun column purification using Sephadex G-50

(Maniatis *et al*, 1982) then stored at -15°C. Specific activity was 6.0 X 10^8 - 2.2 X 10^9 cpm/µg TCA-precipitated DNA.

Transfection Procedure

HepG2 cells were heavily seeded (1/3 splits) in 25cm^2 flasks and grown to semi-confluence. Cells were washed three times with 4ml MEM (no serum), then incubated at 37°C for 1h in the presence of a further 4ml MEM, which was subsequently removed. Flasks were placed on ice and 2ml MEM at 5°C added together with I-[A-CDI] or control protein as indicated in the legend to Figure 5.7. After overnight incubation at 10°C, flasks were again placed on ice, the medium removed, and 1ml MEM at 5°C added together with expression vector as indicated. Flasks were transferred from ice to incubation at 10°C (30min), 20°C (30min), then 37°C (5.0h). At this stage 3ml MEM+S (plus serum) was added and incubation continued for 24h. Flasks were trypsinised and duplicate subcultures established containing 10^6 cells per flask. Growth was continued for a further 24h before the first addition of G418.

Stable HepG2 Expression of neo: G418 Selection

Following the trypsinisation of post-transfection cells, G418 stock solution (24mg total antibiotic/ml H_2O adjusted to pH 7.4 with NaOH and filter sterilised) was added to the medium to give a final concentration of 1200µg/ml (active drug content = approximately 50% of this) and incubation continued for 5 days, after which the medium was changed and fresh G418 added at a final concentration of 600µg/ml. Medium changes were effected at approximately 5 day intervals thereafter, the G418 level being lowered further to and maintained at 500µg/ml. Following the unambiguous appearance of resistant clones (14-21 days) and the death of most non-resistant cells (21-28 days), one set of flasks was stained for the recording of clone distribution. Surviving cells in the other set were trypsinised for subculture and propagation in a G418 maintenance regime and ultimately

subjected to DNA extraction.

Photomicrography of Cells and Clones

HepG2 cells were photographed *in situ* using Ilford FP4 black and white film (ASA 125) in a Nikon F301 single lens reflex camera body attached to a Nikon TMS inverted microscope. Cells were photographed under phase contrast illumination in the presence of a green filter using a X10 objective. At maximum light intensity under these conditions an exposure of 1 or 2 seconds was required.

Macroscopic Visualisation of Clones

Flasks containing clones to be stained were treated as follows: medium was removed, cells rinsed twice with PBS (5ml), fixed with redistilled methanol (5ml, 2min), stained with a 1/10 aqueous dilution of standard Giemsa (5ml, 2-5min), then rinsed generously with water several times. Flasks were drained and air-dried. Stained flasks were photographed in a conventional manner under a combination of incident and transmitted white light using high contrast black and white film, usually Kodak technical pan 2415.

Extraction of Genomic DNA from HepG2 Cells

DNA extraction was by a modification of the method of Shih and Weinberg (1982). Duplicate 75cm^2 flasks of cells were used for each DNA preparation. Each culture flask was decanted, washed twice with 15ml cold PBS, drained, and then gently agitated for 10min in the presence of 3ml lysis buffer (0.5% SDS, 100mM NaCl, 40mM Tris-HCl, 20mM EDTA, pH 7.0). Proteinase K (0.5ml, 2mg/ml, freshly prepared) was added to the total pooled lysate (6ml) to give a final enzyme concentration of 155µg/ml, and the digestion mixture incubated at 37°C for 3.5h. The DNA solution was then deproteinised by means of one 90% (aqueous) phenol extraction followed by two chloroform: isoamylalcohol (24:1v/v)

extractions. DNA was subsequently ethanol precipitated, spooled out on a glass rod, and dissolved overnight in 1ml 0.1SSC at 5°C. Final DNA solutions were dialysed exhaustively against 0.1SSC at 5°C then stored at -15°C. OD_{260}/OD_{280} ratios: 1.73 - 1.93. Yields: 113 - 544µg per extraction (150cm^2 area of almost confluent cells).

Detection of the Stabilised neo Sequence in Transfected HepG2 Genomic DNA

This was carried out by blotting and hybridisation essentially according to Southern (1975). Genomic DNA samples as indicated in the legend to Figure 5.8 were separated in 0.9 - 1.2% agarose gels at neutral pH using either a BioRad Mini Sub DNA cell system (gel size 100 X 65 X 3mm; electrophoresis at 40-50V for 2-3h) or a water-cooled FMC Resolute chamber (gel size 200 X 150 X 6mm; electrophoresis at 40V for 10h). Loading and running buffers were used as described earlier in this section for plasmid gels. Immediately following electrophoresis, gels were denatured without prior depurination (1.5M NaCl, 0.5M NaOH), neutralised (1.5M NaCl, 0.5M Tris-HCl, 1mM EDTA, pH 8.0), then set up for capillary blotting on to nitrocellulose (Amersham Hybond C-extra) using 10 X SSC (1.5M NaCl, 0.15M Na_3 citrate) as blotting medium. Blotting was continued for 12 - 16h. The blotted nitrocellulose membrane was washed (6 X SSC, 5min), drained, and allowed to air dry, after which it was fixed by heat treatment at 80°C (no vacuum) overnight. The fixed blot was stored desiccated at 20°C until use in hybridisation, when it was prewetted with 2 X SSC. The prehybridisation and hybridisation procedures were each conducted in the presence of 50%v/v formamide at 42°C in heat-sealed, heavy-duty polyethylene pockets. Prehybridisation time, 3h; hybridisation time, 20h.

Prehybridisation solution: 5 X Denhardt's solution; 5 X SSC; 50mM sodium phosphate buffer, pH 7.0; 1%w/v glycine; 0.1%w/v SDS; 50%v/v deionised formamide; 500µg/ml freshly boiled carrier DNA. Volume per small gel blot, 10ml; per large gel blot, 20ml.

Hybridisation solution: 1 X Denhardt's solution; 5 X SSC; 20mM sodium phosphate buffer, pH 7.0; 10%w/v dextran sulphate (mol.wt. 500000); 0.1%w/v SDS; 50%v/v formamide; 500µg/ml freshly boiled carrier DNA; 10 - 20ng/ml freshly boiled [^{32}P]-labelled *neo* probe. Volume per small gel blot, 5ml; per large gel blot, 7.5 - 10ml.

Denhardt's solution (X 100): 2%w/v polyvinylpyrrolidone (mol.wt. 360000); 2%w/v ficoll (mol.wt. 400000); 2%w/v BSA.

The hybridised blot was washed under conditions of high stringency as follows and then air dried:

Wash 1 - 2 X SSC, 0.5%w/v SDS, 20°C, 5min;
Wash 2 - 2 X SSC, 0.1%w/v SDS, 20°C, 15min;
Wash 3 - 0.1 X SSC, 0.5%w/v SDS, 68°C, 40min;
Wash 4 - 2 X SSC, no SDS, 20°C, 5 min;
Wash 5 - as wash 4.

Autoradiographic visualisation of the probed blot was achieved by exposing it to X-ray film (Amersham Hyper-film MP) sandwiched between intensifying screens at 20°C. Exposure varied from 10h to several days according to the specific activity of the probe at the time of use. For X-ray film processing, Agfa-Gevaert G150c developer and G334c rapid fixer were employed. For subsequent conventional photography of the autoradiograph, a high contrast black and white film was used, usually Kodak technical pan 2415.

5.3 RESULTS

Physical State of the neo Expression Vectors

Caesium chloride-ethidium bromide density gradient purification of each plasmid vector gave rise to two distinct bands; a less

dense upper band representing the nicked circular form, which allows extensive EB intercalation, and a denser lower band representing the supercoiled form which places considerable constraint upon EB intercalation (Radloff *et al*, 1967) (Figure 5.4A). After withdrawal of the bands followed by removal of the ethidium bromide content by organic extraction and final exhaustive dialysis (section 5.2.2), analysis of the plasmid fractions by agarose gel electrophoresis showed that while the upper band consisted entirely of the nicked circular form, the lower band usually contained plasmid in both the nicked circular and supercoiled configurations (Figure 5.4B). This could have been the result of either inter-band contamination during removal from the gradient, or the nicking of a proportion of the supercoiled population during the subsequent purification steps. The former idea was never supported by observations of reciprocal contamination: the upper band material appeared always to be 'clean'; thus it seemed most likely that post-gradient nicking was indeed taking place, despite routine pretreatment of dialysis tubing (section 2.2.2) and sterilisation of vessels. Furthermore, in the case of pAL-8, the lower band material also contained linearised plasmid (Figure 5.4C), which if present during CsCl separation would have appeared in the upper band. This observation suggests that the treatments following density gradient separation exerted considerable shear forces which, for a large plasmid like pAL-8, were sufficient to physically fragment the DNA double helix. The step during which this shearing might have taken place is that of butanol extraction. In the case of both ptkNEO and pAL-8, the lower band material containing approximately 50% supercoiled plasmid was the fraction used for transfection work.

In confirmation of the known sizes of the two expression vectors, ptkNEO (4913bp) ran rather more slowly than pBR322 (4363bp) during agarose gel electrophoresis, while pAL-8 (>11000bp) ran very much more slowly than either pBR322 or ptkNEO (Figure 5.4C).

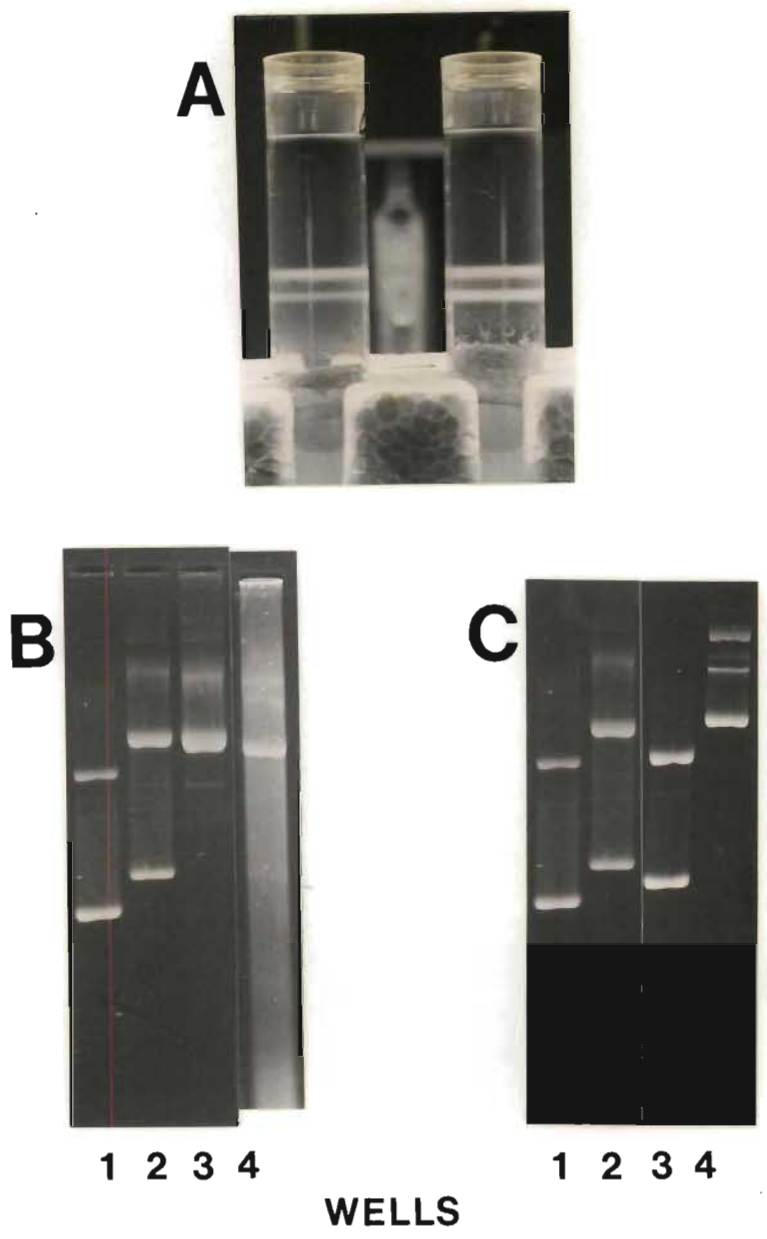


Figure 5.4 Physical state of purified expression vectors

A. Typical condition of vector after CsCl-EB density gradient purification.
Upper band: relaxed (nicked) circular configuration.
Lower band: supercoiled configuration.

B. Content of ptkNEO CsCl-EB gradient bands, after EB removal and dialysis, as shown by agarose gel electrophoresis.
1, pBR322 marker; 2, ptkNEO, lower band material; 3, ptkNEO, upper band material; 4, ptkNEO, crude extract prior to CsCl-EB purification.

C. Relative vector size, as shown by agarose gel electrophoresis.
1, pBR322 (4363bp); 2, ptkNEO (4913bp); 3, pBR322 (4363bp); 4, pAL-8 (>11000bp).

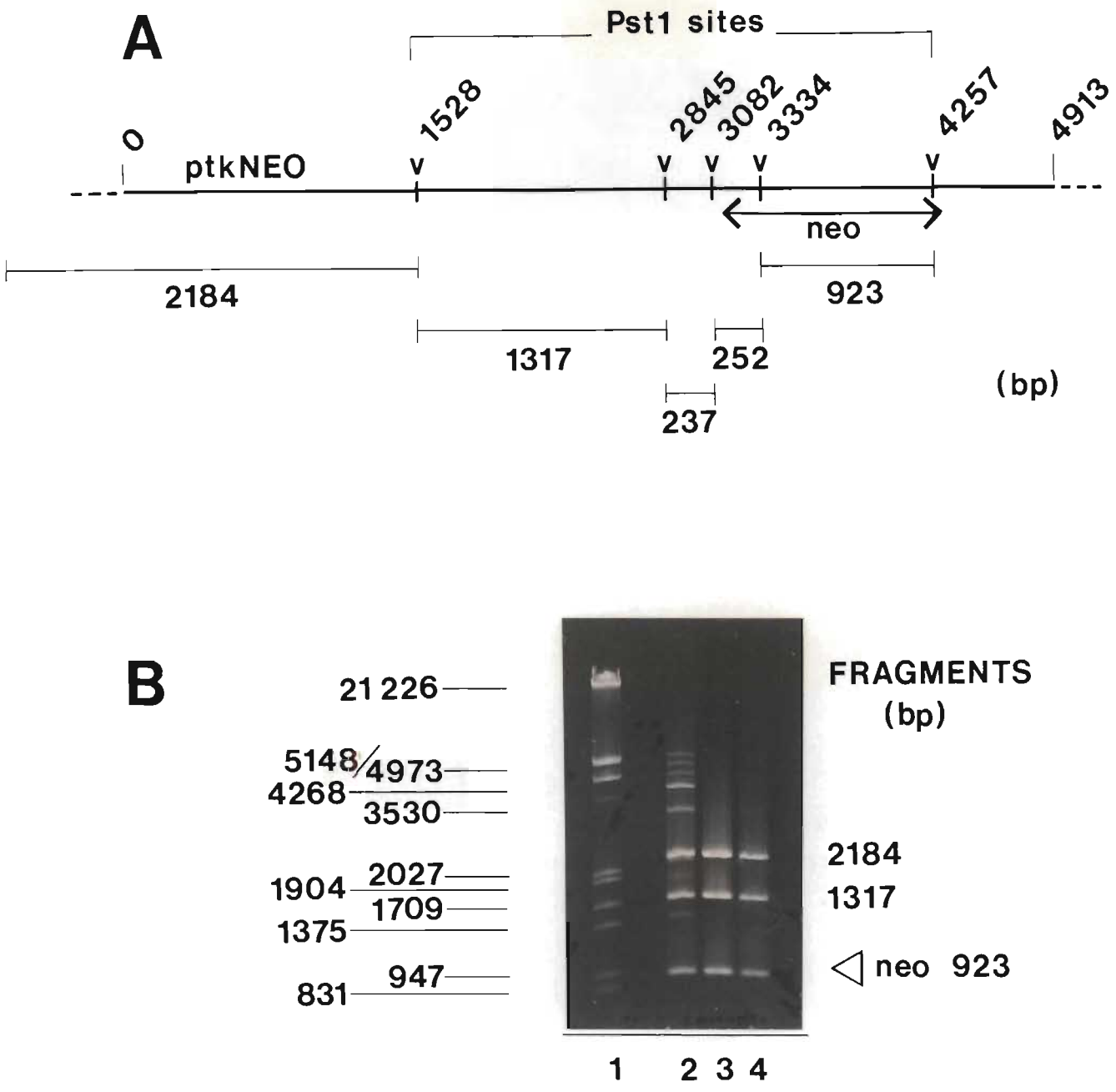


Figure 5.5 Scission of ptkNEO by the restriction endonuclease PstI

A. Cleavage map showing the 5 expected fragments, including a 923bp sequence cut within the *neo* gene.

B. Agarose gel electrophoresis of PstI-restricted ptkNEO. 1, λ-EcoRI-HindIII molecular weight markers; 2-4, PstI digests of 3 independent ptkNEO preparations.

The ptkNEO-PstI Fragment Representing the neo Sequence

The restriction endonuclease PstI, which cleaves DNA at the sequence CTGCA/G, cuts the expression vector ptkNEO in five places, thereby giving rise to five fragments of the following sizes: 2184, 1317, 923, 252 and 237bp. A ptkNEO-PstI cleavage map, based on the complete ptkNEO sequence, is shown in Figure 5.5A, and the fragments, separated on an agarose gel, are depicted in Figure 5.5B. The 923bp fragment, highlighted in Figure 5.5B, consists of the majority of the *neo* sequence, cut within the gene on both sides, thus excluding the flanking HSV1 *tk* sequences altogether. Preparative gels of low melting temperature agarose, in which several identical ptkNEO-PstI digests were run in adjacent wells, were used for the separation and excision of this fragment prior to [^{32}P]-labelling and subsequent use as a probe in Southern hybridisation as described in section 5.2.2.

Detection of the G418 Resistant Phenotype in HepG2 Cells

The course of action of the antibiotic G418 on HepG2 cells is shown in Figure 5.6. The decline and death of the normal, untransfected, G418-sensitive HepG2 cell line over a period of 21 days is depicted in Figure 5.6A. The development, over the same period, of isolated clones of surviving, G418-resistant cells within a *neo*-transfected population is shown in Figure 5.6B. Semiconfluent islands of G418-resistant cells, derived from clone populations by trypsinisation and subculture in G418 medium through several passages, representing a total of over 40 days of G418 treatment, are illustrated in Figure 5.6C.

Figure 5.7 illustrates the final result of a transfection experiment in which clonal colonies have arisen over a period of 28 days in the presence of G418, indicating reception of the *neo* gene and its stable expression. The presence of numerous clones in the experimental flasks E1 - E4 together with the absence of surviving cells in control flasks C1 - C5 provides phenotypic

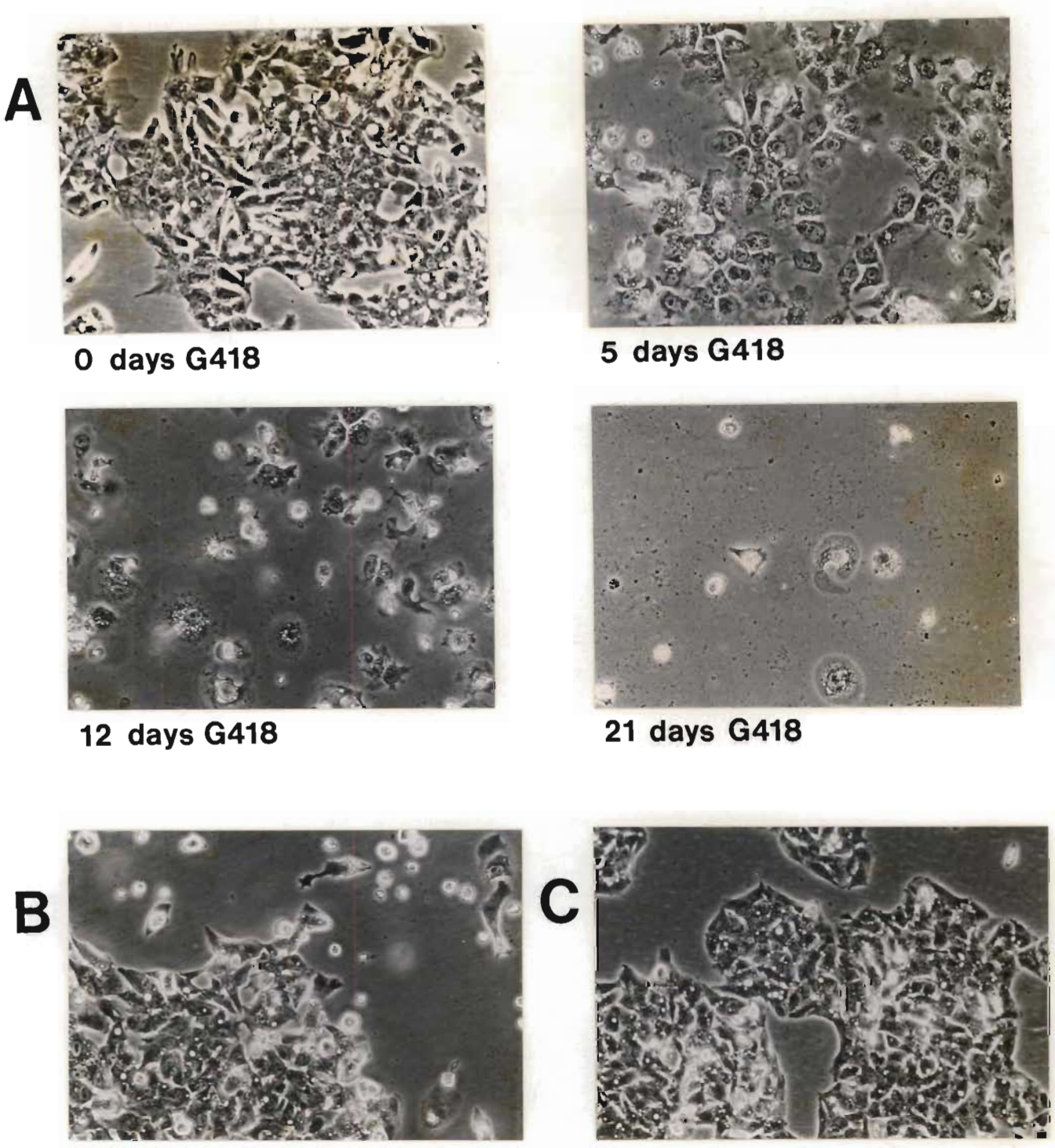


Figure 5.6 Typically observed microscopic effect of the antibiotic G418 on HepG2 survival

Cells were subjected to transfection and/or G418 treatment and photomicrography as described in section 5.2.2.

A. Decline and death of a normal, untransfected cell population.

B. Development of surviving clonal colonies of G418-resistant cells in a pAL-8 transfected cell population.

C. Appearance of a G418-resistant culture after trypsinisation of clonal colony cells from B. and continued growth through several passages in G418 maintenance medium.

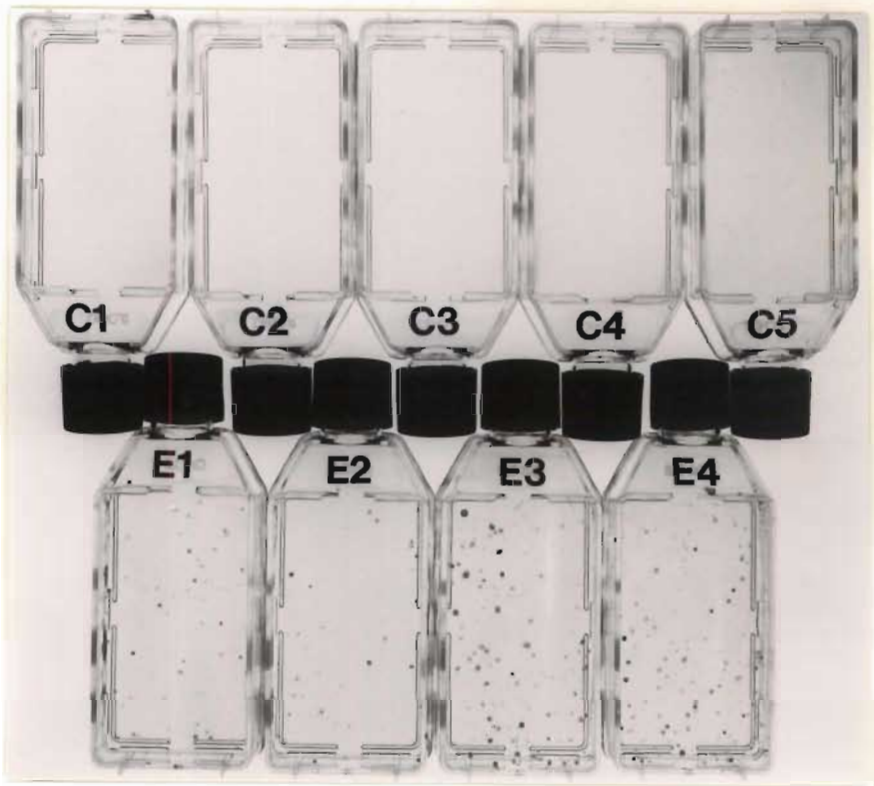


Figure 5.7 Macroscopic appearance of clones resistant to G418 following insulin-[*N*-acylurea albumin] mediated transfer of the expression vectors ptkNEO and pAL-8 to HepG2 cells

Cells were subjected to transfection and G418 selection as described in section 5.2.2. After 28 days of G418 treatment, one complete experimental set of flasks was stained for the visualisation of clones as described in section 5.2.2.

TRANSFECTION FLASKS: E1 and E2, I-[A-CDI] (40µg) plus ptkNEO (4µg); E3 and E4, I-[A-CDI] (40µg) plus pAL-8 (4µg).

TRANSFECTION CONTROL FLASKS: C1, no protein, no vector; C2, no protein, pAL-8 (4µg); C3, unconjugated I (18.25µg) and A (21.75µg) plus pAL-8 (4µg); C4, unconjugated I (18.25µg) and A-CDI (21.75µg) plus pAL-8 (4µg); C5, I-A (40µg) plus pAL-8 (4µg).

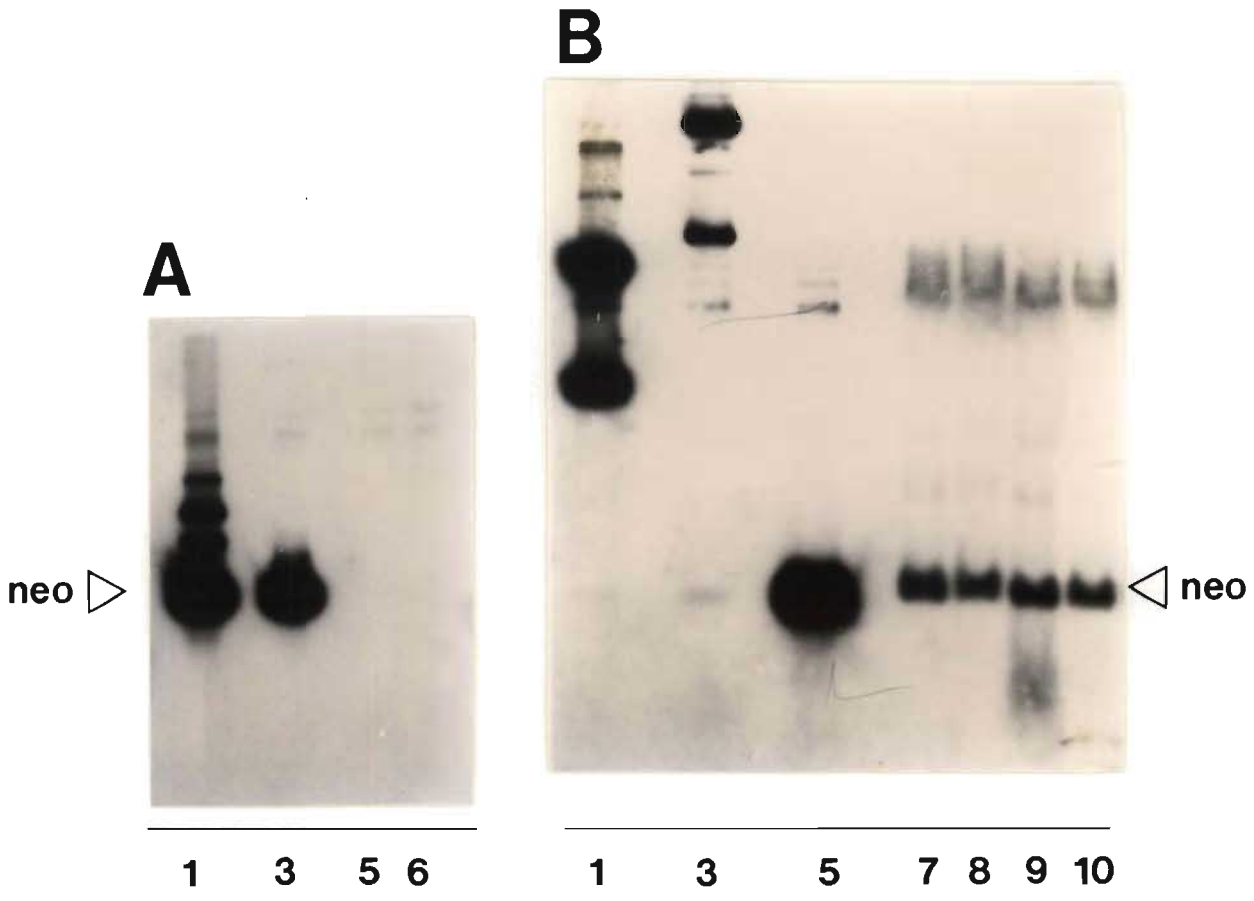


Figure 5.8 Genomic Southern hybridisation: detection of the *neo* sequence in stably G418 resistant HepG2 cells following insulin-[*N*-acylurea albumin] mediated transfection

Cells were subjected to transfection and G418 selection as described in section 5.2.2. After 28 days of G418 treatment, one complete experimental set of flasks was stained for the visualisation of clones (Figure 5.7). Flasks of the duplicate set containing surviving cells (E1 - E4) were trypsinised and each of the cell populations propagated on a G418 maintenance regime, through several passages, for a further 16 days (Figure 5.6C) until there were sufficient cells for DNA extraction. DNA was analysed for the *neo* sequence using as probe a 923bp PstI fragment of ptkNEO within the *neo* gene (Figure 5.5).

A. MINI GEL BLOT : 1, PstI digested ptkNEO (2ng) plus salmon sperm carrier DNA (3µg); 3, *neo* fragment (923bp) from PstI digested ptkNEO (2ng) plus carrier DNA (3µg); 5, PstI digested HepG2 cell line DNA (12µg); 6, PstI digested HeLa cell line DNA (12µg).

B. LARGE GEL BLOT: 1, undigested ptkNEO (4913bp) (1ng) plus carrier DNA (7µg); 3, undigested pAL-8 (>11000bp) (1ng) plus carrier DNA (7µg); 5, *neo* fragment excised from PstI digested ptkNEO (923bp) (1ng) plus carrier DNA (7µg); 7-10, PstI digested DNAs from G418 resistant HepG2 cell populations E1 - E4 respectively (28µg per well).

evidence that DNA transfer takes place only by means of I-[A-CDI] and not as a result of any effects of I-A or unconjugated protein constituents. Frequencies of stable transfection resulting from ptkNEO transfer (mean = 2.0 X 10^{-5}) are lower than for pAL-8 (mean = 5.5 X 10^{-5}). The relationship of these frequencies to the frequencies of stable transfection resulting from other methods of gene transfer is shown in Table 5.3.

Detection of the Stabilised neo Sequence in Transfected HepG2 Genomic DNA

Genotypic evidence for *neo* transfer and its subsequent stabilisation as a replicative entity is provided in Figure 5.8, which shows the results of Southern hybridisation of genomic DNA samples extracted from the HepG2 cell line and from transfected populations derived from it. Figure 5.8A illustrates the absence of *neo* in the DNA of two mammalian cell lines including HepG2. In the case of the HepG2 cells, the DNA required was necessarily extracted prior to transfection experiments, since control cell populations within the experiments themselves ultimately died as a result of G418 treatment (Figure 5.7). Figure 5.8B shows the presence of the *neo* sequence in the transfected HepG2 populations E1 - E4, grown on G418 for a total of 44 days, through several passages, from duplicate flasks of those shown in Figure 5.7.

5.4 SUMMARY AND DISCUSSION

The capacity of insulin-[*N*-acylurea albumin] to carry DNA into mammalian cells bearing insulin receptors has been investigated using expression vectors incorporating the reporter gene *neo* under the control of eukaryotic regulatory elements. The presence of *neo* in recipient cells was detected by means of their survival in the presence of a lethal dose of the

antibiotic G418. Since this system of detection involves long term selection, it provides an indication of not only gene transfer but also replicative maintenance of the gene through many cell divisions and its stable expression. A clear demonstration of the stable expression of the *neo* gene following ptkNEO and pAL-8 transfer to HepG2 cells is provided by (i) the appearance of distinct clones among I-[A-CDI]-vector treated cells grown for 2 - 4 weeks on a G418 regime lethal to untreated cells (Figures 5.6B and 5.7), and (ii) identification of the exogenous *neo* sequence in DNA extracted after clone populations produced in parallel to those shown in Figure 5.7 had been grown for 44 days on the antibiotic (Figures 5.6C and 5.8). Complete lack of survivors in control flasks in which cells were provided with vector in the presence of no protein, I + A, I + A-CDI and I-A respectively (Figure 5.7) constitutes evidence that the gene transfer events depended on an agent capable of binding to both DNA and insulin receptor. Furthermore, free A-CDI, present at a concentration equivalent to that in the I-[A-CDI] conjugate, does not bring about transformation, indicating that non-specific cation-induced uptake is not involved. These results suggest that gene transfer has taken place under the direction of insulin via the endocytotic pathway.

Observations made previously in relation to the solubility properties of the insulin-[*N*-acylurea albumin]-DNA complex had suggested that a degree of insolubility might be encountered under transfection conditions (section 3.4). In the transfection protocol, therefore, care was taken to avoid the possibility of insoluble material being precipitated in the medium. This was achieved by establishing conjugate-receptor binding during a pre-incubation period, following this with removal of unbound conjugate, and only then adding expression vector DNA for attachment to the receptor-bound carrier. Receptor cycling was inhibited by temperature control until the DNA binding step was complete. In addition to circumventing any insolubility problems, this approach to binding allowed the addition of excess amounts of conjugate and DNA at each step,

maximised binding and possibly increased transfection efficiency. The alternative would have been to ensure removal of particulate material by membrane filtration of the complex before addition to cells. This might prove to be a suitable approach for future *in vivo* gene transfer studies, but further investigation into the interactions of soluble I-[A-DI]-DNA complex and components of the medium would be a desirable prerequisite.

During the period of initiation and development of the transfection and selection protocol, no precedent could be found in the literature for the use of the *neo*-G418 system in HepG2 cells. Suitable concentrations of G418 for the killing of untransfected cells had to be determined empirically, therefore, using regimes applicable to other cell lines as guides. Gorman (1985) has pointed out the importance of this determination, commenting that cells may escape the lethal effect if the dose of antibiotic is too low, and suggesting that initial trials should employ G418 concentrations ranging from 200 - 1000µg/ml. Examples of concentrations found suitable for use with mammalian cell lines include 250µg/ml (active drug) for murine teratocarcinoma cell lines (Jakob and Nicolas, 1987), 300µg/ml (active drug) for transformed monkey cell lines (Gerard and Gluzman, 1985), 500µg/ml (whole drug) for a wide variety of mammalian cell lines (Flyer *et al*, 1983), 625µg/ml (whole drug) for NIH 3T3 cells (Anklesaria *et al*, 1987), and 1000µg/ml (whole drug) for NIH 3T3 cells (Cepko *et al*, 1984). The active drug component of most commercial preparations of G418 is in the region of 40 - 60%. In the G418 used for the work described here it ranged from 48.6 - 53.0% according to production lot, and was regarded as 50% for convenience. The antibiotic regimen eventually found suitable for HepG2 cells consisted of an initial G418 treatment at 600µg/ml (active drug) for 5 days, a further treatment at 300µg/ml (active drug) for 5 days, followed by maintenance treatment at 250µg/ml (active drug), medium being changed at 5 day intervals until the selection process was complete. The G418 sensitivity of HepG2 cells thus appears to

be intermediate in the range shown by other cell lines. However, it was notable that while the G418-induced death of normal HepG2 cell populations was fairly advanced after 10 days or so, it was not fully complete until 3 - 4 weeks after the start of treatment. Increasing the dosage in the maintenance medium made little significant difference to the temporal factor. The killing period quoted for other cell lines is generally shorter: 1 - 2 weeks (Gorman, 1985; Flyer *et al*, 1983). It is interesting to consider the possibility that a highly differentiated liver cell line such as HepG2 might retain the normal hepatic ability to detoxify drugs, including compounds such as the gentamicins. If so, the process of *neo* selection would probably be retarded, even at high concentrations of G418. In normal liver tissue *in vivo*, aminoglycoside antibiotics may be inactivated in the microsomal apparatus by conversion to *N*-glucuronides (A. Papaphilippou, personal communication).

An important factor in interpreting the growth of G418-resistant clonal colonies is that it can be assumed that the medium in which the clones develop maintains a high concentration of unchanged antibiotic. In the majority of cases of naturally occurring resistance studied in prokaryotic systems, aminoglycoside antibiotics are not inactivated or detoxified in the culture medium (Davies and Smith, 1978). It is thought that the mechanism of resistance depends, rather, on intracellular enzymatic modification interfering with or blocking the transport of the drug into and within the cell interior, possibly by direct interaction of the modified antibiotic with a carrier molecule. Little is known about the uptake of aminoglycoside compounds by sensitive cells, however, so the means by which such uptake processes may be inhibited in resistant types remain speculative (Davies and Smith, 1978). The significance to the present work of what is known of the mechanism of resistance is that there is little chance of artefactual clones having arisen as a result of extracellular G418 degradation.

Expression of the *neo* gene following insulin-[*N*-acylurea albumin] transfection has been shown to be extremely stable. The clones which arose in duplicate transfection flasks over 4 weeks were either fixed and stained (Figure 5.7) or trypsinised and grown to high cell numbers through several subsequent subculture passages in G418 medium (Figure 5.6C) for the eventual extraction of genomic DNA. This DNA, utilised in Southern blotting for the detection of the *neo* gene (Figure 5.8), was thus derived from cells which had experienced 44 days of continuous G418 treatment and many cell divisions. In parallel experiments, cells were sometimes maintained in G418 medium for even longer periods: up to 78 days. Recent observations during further transfection experiments, incomplete at the present time, suggest that G418-resistant cells do not lose the *neo* gene even when grown through numerous passages in medium free of the antibiotic, that is, in the absence of selection.

The size of the G418-resistant clones developed (Figure 5.7) varied greatly within each treated population regardless of vector. This phenomenon has measurable parallels in plant genetics: individual *neo*-positive plantlets grown up from Ti plasmid-transformed tissue show wide quantitative difference in resistance to kanamycin; this is generally thought to indicate that expression of inserted genes is enhanced or suppressed according to genomic position, which varies as a result of random integration (Rogers *et al*, 1986). The wide variations in insertional position which appear to be indicated by differences in clonal size do emphasise the point that *populations* of clones are themselves unlikely to be clonal in the strict sense, but rather genetically heterogeneous, although in a manner undetectable by the methodologies used in the work presented here.

In enumerating clones, those which were particularly small were ignored; transformation frequencies were estimated conservatively thereby. Results (Table 5.3) show I-[A-CDI]-

AGENT/MODE OF TRANSFECTION	CELL LINE	REPORTER GENE	TRANSCRIPTION UNIT/ PROMOTER	FREQUENCY OF STABLE TRANSFORMATION (per 10^6 cells)	REFERENCE
Ca phosphate	HepG2	various	various	1	Darlington (1987)
Bacterial Protoplasts	K562	*neo*	SV40 (early)	1 - 10	Tsao *et al* (1987)
I-[A-CDI]	**HepG2**	*neo*	**HSV1 tk**	**20**	**section 5.3**
I-[A-CDI]	**HepG2**	*neo*	**SV40 (early)**	**55**	**section 5.3**
Recombinant retroviruses	various	*neo*	M-MSV	100	Flyer *et al* (1983)
Scrape loading	REF / D2	*neo*	SV40 (early)	186	Fechheimer *et al* (1987)
Polycation and DMSO	chicken embryo	*src*	RSV (whole genome)	up to 800	Kawai and Nishizawa (1984)
Electroporation	mouse L	*neo*	SV40 (early)	1000	Stopper *et al* (1987)
Microinjection	various	various	various	1000 or more	Graessmannn and Graessmann (1983); Celis (1984)

Table 5.3 Stable transformation of mammalian cell lines: efficiency of the insulin-[*N*-acylurea albumin] mode of gene transfer in relation to other methods

induced stable gene transfer to be approximately 20 times (ptkNEO) and 55 times (pAL-8) more efficient than typical stable transformation of HepG2 cells brought about by the calcium phosphate procedure (Darlington, 1987), although only 0.2 (ptkNEO) and 0.5 (pAL-8) the efficiency of *neo* transfer to mammalian cells effected via retroviral constructs (Flyer *et al*, 1983).

The notable difference in frequency between ptkNEO- and pAL-8-induced expression of *neo* might indicate that SV40 promoter sequences signal transcription in HepG2 cells more effectively than do the HSV1 *tk* control regions. The inclusion of the tandem 72bp repeat enhancer in the SV40 early promoter region upstream of *neo* in pAL-8 constitutes one possible explanation for the observed difference. The dramatic effect of this enhancer both in SV40 and in other chimaeric recombinants, independent of position and orientation, has been well documented (Moreau *et al*, 1981; Laimins *et al*, 1982). Furthermore, Capecchi (1980) has shown specifically that its presence increases the transformation efficiency of the HSV1 *tk* gene.

An additional or alternative explanation may lie in the fact that pAL-8 contains the origin of replication of SV40, which signifies the possibility of the transfected vector existing episomally in the cytoplasm in a stable manner, transmissible to daughter cells during division. For this type of replication to take place, only one virally encoded protein is required: the T antigen (Ishimi *et al*, 1988). In transfection studies with transformed monkey cells, Gerard and Gluzman (1985) showed that certain recombinant vectors carrying the *neo* sequence and the SV40 origin of replication, specifically designed for cytoplasmic autonomy, were maintained in an episomal form in recipient cells, persisting even in the absence of biochemical selection. In their work, evidence for episomal copies of *neo* was provided by Southern hybridisation of low molecular weight DNA extracted by the Hirt procedure (Hirt, 1967); chromosomal

DNA extracts were not tested. However, in a rigorous investigation involving the introduction of various conventional pSVneo plasmid vectors (including pSV_2neo) into mammalian cells, Southern and Berg (1982) showed clearly that during the course of G418 selection the *neo* gene became exclusively associated with high molecular weight (genomic) DNA, autonomously replicating forms being absent as indicated by complete lack of the *neo* sequence in the Hirt fraction. At the present time it is not known whether the T antigen gene is incorporated into the SV40 regions of the pAL-8 construct, so the possibility of episomal replication of *neo* cannot be excluded. However, it is probable, as discussed in section 5.1, that the pAL-8 SV40 components are closely related to those in the pSV_2 vectors, and that the stable expression of *neo* observed is solely the result of genomic integration. That chromosomal integration events have occurred is supported by the fact that *neo* is identifiable in transfected genomic DNA (Figure 5.8B) and by the demonstration of stable expression following transfection with ptkNEO, which carries no eukaryotic origin of replication.

The design of the insulin-[*N*-acylurea albumin] gene carrier is such that its most important role is played out at the plasma membrane in effecting transportation of extracellular DNA to the cell interior. The ligand-directed passage of DNA across the membrane is likely to be highly efficient. During subsequent intracellular processing, however, the route of DNA is less controlled, and numerous hazards in the form of enzymatic or membrane barriers must be encountered. Ultimate delivery to the nucleus would seem to be a rare event. The use of insulin as ligand might offer an advantage in this respect. In contrast with some ligand proteins which are delivered via membrane-bound vesicles to lysosomes, it is increasingly evident that insulin degradation inside the cell takes place in endosomal rather than lysosomal compartments (Sonne, 1988). Thus DNA still associated with the I-[A-CDI] complex after entry may not be directed actively towards lysosomal nucleases. It is tempting to speculate further upon the possibility of insulin-[*N*-acylurea

albumin] directing passage of vector DNA into the nucleus via the insulin receptors situated in the nuclear membrane (Vigneri *et al*, 1978). Attractive as this idea is, DNA binding studies have shown that in the ionic environment of the cytoplasm the I-[A-CDI]-DNA complex is likely to be at least partially dissociated (Figure 3.5B). A proportion of vectors reaching the cytoplasm are probably, therefore, free. The reversibility of binding is a desirable feature in general, since the DNA released is completely unmodified. However, if a soluble insulin carrier with a higher affinity for DNA could be devised, the advantage of dissociability would not be lost, and the postulation of insulin-directed endosomal transport and nuclear transfer might be a realistic one.

CHAPTER SIX

CONCLUDING REMARKS

6.1 THE WORK IN PERSPECTIVE

The carbodiimide modification of proteins to produce *N*-acylurea derivatives with DNA-binding properties and their potential in gene transfer has been reported previously (Huckett *et al*, 1986; Ariatti and Hawtrey, 1987). In this thesis it has been demonstrated that cross-linking insulin to carbodiimide-modified albumin produces a soluble macromolecule capable of both specific recognition of the insulin receptor and DNA transport, thereby affording a means of targeted transfection from solution without the involvement of particulate intermediates, excessive foreign chemicals, cell disruption or infective agents. Evidence is provided which suggests that by means of this DNA carrier system, in a manner consistent with the proposed model (Figure 1.1), the exogenous gene *neo* from the bacterial transposon Tn5 has been inserted through the insulin receptor endocytotic pathway into the HepG2 hepatoma cell line, integrated into the genome, and stably expressed.

Exploitation of the efficiency and specificity of cell entry afforded by receptor-mediated endocytosis had been suggested in various ways prior to this work being undertaken. The targeting of drugs, enzymes and genetic material had been considered by various research groups in this context, as discussed briefly in section 1.3.5. For example, Poznansky and coworkers achieved delivery of α-1,4-glucosidase to muscle cells and hepatocytes using a covalently conjugated enzyme-albumin-insulin complex (Poznansky *et al*, 1984)), and Pastan's group (Cheng *et al*, 1983) devised a mechanism of covalent attachment between nucleic acid

and protein with the purpose in view of genetic correction via endocytosis. More recently, an alternative methodology for the preparation of oligonucleotide-peptide hybrids has been reported with the stated aim of targeting anti-sense oligonucleotides into cells via endocytosis for the therapeutic inhibition of gene expression (Haralambidis *et al*, 1989). In contrast with applications of endocytotic transport such as these, the method of gene transfer expounded in the present work combines the elegance of ligand-directed carrier insertion with non-covalent reversible binding of the DNA to the carrier, a possible advantage in the cell interior. Work by Wu and Wu published during the course of this investigation illustrates use of a parallel approach: by employing as carrier a conjugate of asialoorosomucoid and high molecular weight cationic poly-L-lysine, which binds DNA electrostatically in much the same manner as *N*-acylurea protein, those workers have demonstrated transient expression of the bacterial *cat* gene in HepG2 hepatoma cells in culture (Wu and Wu, 1987; Wu and Wu, 1988a) and in rat liver *in vivo* (Wu and Wu, 1988b). The work described in this thesis adds, therefore, to a growing body of evidence that receptor mediated endocytosis provides an effective pathway by which exogenous molecules may be directed into the interior of target cells, and in particular shows that highly efficient stable genetic transformation may result from the use of this facility.

The fact that the work undertaken in the present investigation is so closely allied to the work reported by Wu and Wu (1987, 1988a) gives the latter an important status in relation to the thesis: comparisons are especially pertinent. For example, the poly-L-lysine used by Wu and Wu is of the same molecular size (mean M_r = 69000) as the *N*-acylurea albumin used here (m wt = 68000 for the monomer). However, conjugation of the poly-L-lysine to the ligand asialoorosomucoid (ASOR) was in a 1:5 mole ratio and employed *N*-succinimidyl 3-(2-pyridyldithio) propionate as coupling agent, compared to conjugation of *N*-acylurea albumin to insulin in a 1:10 ratio by means of

glutaraldehyde as in the present work. A more relevant difference is that Wu and Wu devised a simple but effective method for determining the proportion of ASOR-poly-L-lysine conjugate that should be mixed with plasmid DNA in order to optimise the DNA content of a soluble complex. Thus mixtures of [^{125}I]-labelled conjugate and DNA in 2M NaCl were set up in which the concentrations of conjugate varied (0 - 160nM) but the concentration of DNA remained constant (87nM). Mixtures were incubated for 1h at 25°C, then dialysed against 0.15M NaCl for 24h using low molecular weight cut-off dialysis tubing. Samples were subsequently filtered through membranes with a pore size of 0.2μM to ensure that the complexes did not contain precipitates. Fluorometric measurements of DNA in the filtrates showed that there was no significant loss of DNA due to the filtration step. The samples were then electrophoresed in an agarose gel, after which the DNA was visualised by ethidium bromide staining and the conjugate located by autoradiography. The results of this procedure enabled the authors to identify the 2:1 mole ratio of conjugate to DNA as the one optimising complex formation. This ratio was used, therefore, in all subsequent transfection work; moreover, the complexes were always subjected to a precautionary prefiltration step. In further contrast to the present work, ASOR-poly-L-lysine transfection was carried out in the presence of foetal calf serum and a Ca^{++} supplement (5mM), pSV_2CAT was employed as expression vector, and the SK-Hep 1 line (ASOR receptor-negative) used as a cellular control. Receptor-positive HepG2 cells used as recipients demonstrated transient expression of the *cat* gene at twice the efficiency of the calcium phosphate method in the authors' hands.

While the particular strength of the present work compared to that of Wu and Wu lies in its clear demonstration of stable expression following endocytotic gene transfer (Figures 5.7 and 5.8), there are several aspects of the approach taken by those authors, mentioned above, which could be gainfully introduced into the insulin-[*N*-acylurea albumin] methodology as reported to date. The most important is the measurement and optimisation of

soluble conjugate-DNA complex. Of lesser importance, but of considerable value in confirming the mechanism of gene transfer, is the use of an endocytosis-negative cell line as control. Of some further interest would be an investigation into the effect of Ca^{++} on endocytotic transfection.

6.2 PROSPECTS FOR DEVELOPMENT OF THE WORK

Various extensions to the studies contributing to this thesis have suggested themselves during the course of the work. Together with the modifications suggested by the approach used by Wu and Wu, they fall into six categories as follows:

Mechanism of Conjugate-DNA Binding

The manner in which insulin-[*N*-acylurea albumin] conjugate binds DNA non-specifically is presented and discussed in chapter 3. The residual DNA binding shown by the conjugate at high ionic strength (Figure 3.5B) is intriguing, and deserves some further attention. In particular, two of the possible explanations offered in section 3.4 could be investigated. Firstly, non-electrostatic interaction in the form of hydrogen bonding, suggested by Ariatti and Hawtrey (1987) to have particular potential in relation to guanine residues in single-stranded regions of DNA, could be studied by examining the binding of I-[A-CDI] to single-stranded DNAs such as M13 DNA and single-stranded poly-G, or by employing oligo-G fragments as competitors of double-stranded DNA binding. Secondly, the possibility that the residual binding is artefactual due to partial precipitation of the I-[A-CDI]-DNA complex could be examined by introducing a prefiltration step into the procedure. This latter point of investigation would be covered should the conjugate-DNA complex optimisation protocol of Wu and Wu (1987; 1988a), described above, be adopted.

Specificity of Conjugate-Receptor Binding in Transfection

The binding reaction between I-[A-CDI] and HepG2 cells is described and discussed in chapter 4. On the basis that the binding kinetics (Figure 4.2) indicate clearly that the conjugate interacts specifically with the insulin receptor and not with non-specific sites on the cell surface, the transfection work which follows, presented in chapter 5, is assumed to depend upon receptor-determined endocytotic uptake. Further evidence that the uptake is indeed a receptor specified process is provided by the apparent inability of A-CDI to mediate gene transfer (Figure 5.7). However, as discussed above in relation to the work of Wu and Wu (1987;1988a), additional support for the specificity of the process could be provided by the use in transfection of control cells lacking the appropriate receptor. Few cells, if any, demonstrate complete absence of the insulin receptor (Rosen, 1987). Convenient controls would best be selected, therefore, from mutant cell lines in which insulin receptor function is seriously impaired; alternatively, IM-9 lymphocytes, which carry insulin receptors but do not exhibit endocytosis, could possibly be used for this purpose.

Parameters of Transfection

A substantial and meaningful extension to the work would be to characterise the important parameters of transfection using the insulin receptor pathway, and to establish optimal conditions for both gene transfer and subsequent expression. For a study of this nature, it would be prudent to use, in the initial stages, a transient expression assay to provide a rapid and quantitative indication of transfection and cytoplasmic expression. A well tested transient expression vector such as pSV_2CAT would be convenient for this purpose, although it would also be possible to continue using ptkNEO and pAL-8 and to assay for the transient expression of *neo* using techniques for the identification and quantification of APH(3')II enzyme, electrophoretically separated from a crude cell extract, such as

the method of Reiss *et al* (1984) or a Western blotting technique. The importance to expression of such factors as time after transfection, culture medium composition (including Ca^{++} concentration), state of the recipient cells, the optimal concentrations of carrier and donor DNA, and the physical state of the DNA could then be investigated with relative ease. It has been noted by a number of workers that the expression of transfected DNA depends upon its topology, although the character of the dependency varies from system to system. In CV-1 and L cells, supercoiled DNAs yield markedly higher levels of expression than their linerarised equivalents (Weintraub *et al*, 1986); in plant protoplasts, however, linear forms of plasmid DNA have been observed to be superior to supercoiled structures as active templates for gene expression (Ballas *et al*, 1988). Further investigations, possibly using a wider range of expression vectors, could lead to reliable and reproducible mean values for the frequencies of transient and stable expression, thereby establishing the relationship between the two. In the case of vectors based on SV40 regulatory sequences it would be important to extend the studies of stable expression to include Southern blotting of both a Hirt DNA fraction (Hirt, 1967) and a genomic DNA extract in order to establish the replicative mode of the inserted plasmid and thus the occurrence of chromosomal versus episomal maintenance of the reporter gene (section 5.4). Establishment of transfection frequencies using the classical calcium phosphate technique with the same range of vectors would provide a valuable general comparative measure of the efficiency of the transfection system in HepG2 cells.

Sophistication of the Method

Methodological improvement is seen to be a logical and desirable direct development of the present work, which suffers from the interrelated limitations of poor conjugate solubility, poor conjugate-DNA complex solubility, and only moderate conjugate-DNA affinity under physiological conditions. Some of these problems are due to the limited solubility of insulin

itself at neutral pH, and might be overcome only by turning to a more soluble alternative ligand. It has already been established by other members of the research group that when readily soluble ligand proteins such as transferrin and asialoorosomucoid (α_1 acid-glycoprotein) are modified with carbodiimides containing a quaternary nitrogen function, thus eliciting a stronger electrostatic interaction with DNA than in the case of tertiary carbodiimide modification (Figure 3.5A), the complexes so formed demonstrate greater flexibility in solution than the equivalently modified insulin-albumin conjugate (Huckett *et al*, 1986; N. Moodley, unpublished results). However, alternative ways of modifying insulin could possibly be found; it is also possible that establishment of optimal conditions for transfection using the tertiary derivative I-[A-CDI], as mentioned in the previous subsection, might in itself resolve sufficient of the difficulties to achieve a desirable degree of technical amendment.

Intracellular Fate of Transfected DNA

A completely separate area of investigation, one which might constitute a project in its own right, is that of following the transport route and survival of donor DNA within the cell interior following transfection. The interaction, if any, of the DNA with lysosomes, endosomal elements and the nuclear membrane would be among the primary objectives. This type of study would necessitate employing the kinds of techniques used to follow the fate of ligands themselves: radio- or immunolabelling of the DNA, and its subsequent intracellular visualisation or observed association with subcellular fractions at various times after transfection, combined with the use of a variety of inhibitors. Work of this nature is fairly difficult, and interpretations even more so. However, any data concerning localisation of the transfected DNA in the various cytoplasmic pathways and its possible occurrence in the nucleus would be intrinsically interesting, and would have additional value in relation to ongoing work concerning the relative frequencies of

transient and stable expression.

Application of the Method in vivo

Ultimately, it would be challenging to undertake the receptor-mediated delivery of a foreign gene to animal tissue *in vivo*; in fact, there are arguments for moving in this direction without delay (section 6.3). The most recently published work of Wu and Wu (1988b) models this approach using the pSV_2CAT system in rats. A similarly simple transient expression vector would be the most reasonable and practical one to use with the *N*-acylurea protein method in the first place, and would lend itself to investigating the organ targeting potential of a specific ligand. Insulin might prove a poor choice of ligand for targeted gene delivery, since its cognate receptor exists in almost every type of cell (Rosen, 1987); however, it is found in greatest concentrations in liver and fat, and it would be of value to determine whether those tissues would constitute prime areas of transfection activity *in vivo*. Since the present work has shown that long term expression of *neo* is possible using an *N*-acylurea protein carrier *in vitro*, it would be particularly interesting to proceed eventually to the use of a suitable stable expression system in a whole animal model.

6.3 SPECULATIONS

The possible avenues of further investigation briefly outlined in section 6.2 are necessarily seen in the context of the current status of molecular biology in general and foreign gene expression in particular. Should it be possible for some of these lines of study to be undertaken in the future, it is clear that their emphasis would be modified in accordance with contemporary findings in the field. While the findings themselves are not predictable, it is possible to speculatively suggest the areas in which the greatest progressive shifts might

occur.

At the present time, gene transfer research appears to be most active in the following four areas:

The production of commercially important transgenic animals and plants. Recent reviews of the genetic engineering of livestock (Pursel *et al*, 1989) and crop plants (Gasser and Fraley, 1989) indicate that dramatic progress is being made in these fields. The expansion of gene transfer practice within the agricultural industry is such that regulatory approval and patent protection are key issues; these in turn are determining, at the highest levels of government, the nature and degree of further investment in the relevant technology.

Therapeutic gene replacement in vivo. Increasingly numerous and sophisticated reports of gene therapy in whole animal models are appearing in the literature. In general, the momentum within this area of endeavour is being generated by the success of the approach involving transplantation of recombinant cells. For example, transformed endothelial cells introduced into denuded iliofemoral arteries of swine have been shown to become implanted into the vessel wall and to express the recombinant gene there 2 - 4 weeks later (Nabel *et al*, 1989). In a study of related significance reported by Mulligan's group, prosthetic vascular grafts seeded with genetically modified endothelial cells have been implanted successfully into dogs, survival of the recombinant cells being demonstrated after 5 weeks (Wilson *et al*, 1989). As a measure of the medical relevance and urgency of this type of work, Friedmannn (1989) has reviewed current progress towards gene therapy in humans, closely following a number of previous reviews (section 1.2.3).

Mammalian gene targeting by homologous recombination. The current status of intrachromosomal gene targeting has been reviewed by one of the major research workers in this area (Capecchi, 1989). Significant accomplishments have been made in

the last two years, especially in terms of the development of methods for the identification in embryonic stem cells of a desired recombinational event. These advances, added to the importance in human gene therapy in particular of correct gene placement, augur well for the future growth of this aspect of genetic transfer.

Gene delivery to specific organelles. A number of recent studies have emphasised an interest in gene transfer to those organelles which contain non-nuclear DNA: mitochondria and chloroplasts. Since a number of crucially important enzymes are encoded in these organellar genomes, gene transfer which is specifically directed to them provides a tool with which fundamental organelle controlled processes may be altered and thereby studied. Using the microprojectile method, for instance, *Chlamydomonas* chloroplast genome modification (Boynton *et al*, 1988) and *Saccharomyces* mitochondrial transformation (Johnston *et al*, 1988) have been achieved, and by means of conjugating a short piece of DNA (24bp) to a mitochondrial precursor protein, Vestweber and Schatz (1989) have demonstrated mitochondrial DNA uptake in yeast. Related to this type of approach is the possibility of introducing DNA into the nucleus in a directed manner via nuclear proteins or nuclear membrane receptors. Kaneda *et al* (1989), for example, have shown that a fivefold increase in expression is observed when DNA is cointroduced with nuclear protein into rat liver.

While the first of the above areas of research activity is unlikely to have serious impact on the future prospects of receptor-mediated gene transfer (and vice-versa), the last three mentioned might have some bearing on its development and application. An effective method of soluble DNA transfer with the potential for targeting to cells or tissues, as modelled in this study and concomitantly by Wu and Wu (1987; 1988a), does fill a gap in the range of available methodologies (Table 1.1), and the evidence that it is applicable *in vivo* (Wu and Wu,

1988b) makes the possibility of its therapeutic application realistic and exciting. In the current research climate, therefore, the transition of the receptor-mediated approach towards whole animal work and its refinement in terms of intracellular targeting would seem to be important. In addition, the endocytotic method of molecular transfer appears to show promise in relation to organellar gene delivery, especially perhaps in the case of insulin-based carriers. The fact that a proportion of endocytosed insulin receptors normally remains undegraded in the cell and that receptors are detectable at sites on the nuclear membrane has been noted in section 1.3.4, and its implications for transport of DNA into the nucleus discussed in section 5.4. For this reason, and especially in the light of the apparent growth of interest in subcellular gene targeting mechanisms, further development of insulin-directed gene transfer in particular would seem to have a modest degree of profitable potential, worthy of pursuit.

LITERATURE CITED

1. Adair, G.M., Nairn, R.S., Wilson, J.H., Seidman, M.M., Brotherman, K.A., MacKinnon, C. and Scheerer, J.B. (1989). Targeted homologous recombination at the endogenous adenine phosphoribosyltransferase locus in Chinese hamster cells. *Proc. Natl. Acad. Sci. USA* 86: 4574-4578.

2. Aden, D.P., Fogel, A., Plotkin, S., Damjanov, I. and Knowles, B.B. (1979). Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. *Nature* 282: 615-616.

3. Anderson, J.E., Ptashne, M. and Harrison, S.C. (1985). A phage repressor-operator complex at 7Å resolution. *Nature* 316: 596-601.

4. Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. (1981). Structure of the cro repressor from bacteriophage λ and its interaction with DNA. *Nature* 290: 754-758.

5. Anklesaria, P., Sakakeeny, M.A., Klassen, V., Rothstein, L., Fitzgerald, T.J., Appel, M., Greenberger, J.S. and Holland, C.A. (1987). Expression of a selectable gene transferred by a retroviral vector to haematopoietic stem cells and stromal cells in murine continuous bone marrow cultures. *Exp. Hematol.* 15: 195-202.

6. Anson, D.S., Hock, R.A., Austen, D., Smith, K.J., Brownlee, G.G., Verma, I.M. and Miller, A.D. (1987). Towards gene therapy for hemophilia B. *Mol. Biol. Med.* 4: 11-20.

7. Ariatti, M. and Hawtrey, A.O. (1987). The binding of DNA to water soluble carbodiimide modified proteins. *Med. Hypoth.* 24: 29-41.

8. Assoian, R.K. and Tager, H.S. (1981). [(^{125}I)Iodotyrosyl^{B1}] insulin. Semisynthesis, receptor binding, and cell-mediated degradation of a B chain-labelled insulin. *J. Biol. Chem.* 256: 4042-4049.

9. Avery, O.T., MacLeod, C.M. & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* 79: 137-158.

10. Ayad, S.R. and Fox, M. (1968). DNA uptake by a mutant strain of lymphoma cells. *Nature* 220: 35-38.

11. Ballas, N., Zakai, N., Friedberg, D. and Loyter, A. (1988). Linear forms of plasmid DNA are superior to supercoiled structures as active templates for gene expression in plant protoplasts. *Plant Mol. Biol.* 11: 517-527.

12. Baltimore, D. (1985). Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome. *Cell* 40: 481-488.

13. Barber, M.A. (1914). The pipette method in the isolation of single microorganisms and in the inoculation of substances into living cells. *Philippine J. Sci.* 9: 307-325.

14. Bauminger, S. and Wilchek, M. (1980). The use of carbodiimides in the preparation of immunizing conjugates. *Methods in Enzymology* 70A; 151-159.

15. Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980). The ovalbumin gene-sequence of putative control regions. *Nuc. Acids Res.* 8: 127-142.

16. Benoit, J., Leroy, P., Vendrely, C. and Vendrely, R. (1957). Des mutations somatiques dirigees sout-elles possibles chez les Oiseaux. *Compt. Rend. Acad. Sci.* 244: 2320-2321.

17. Benoit, J., Leroy, P., Vendrely, R. and Vendrely, C. (1960). *Biochem. Pharmacol.* 4: 181-

18. Berdy, J., Pauncz, H.K., Vajna, Z.M., Horvath, G., Gyimesi, J. and Koczka, I. (1977). Metabolites of gentamicin-producing *Micromonospora* species. I. Isolation and identification of metabolites. *J. Antibiot. (Tokyo)* 30: 945-954.

19. Bergeron, J.J.M., Cruz, J., Khan, M.N. and Posner, B.I. (1985). Uptake of insulin and other ligands into receptor-rich endocytic components of target cells: the endosomal apparatus. *Ann. Rev. Physiol.* 47: 383-403.

20. Bergeron, J.J.M., Sikstrom, R., Hand, A.R. and Posner, B.I. (1979). Binding and uptake of ^{125}I-insulin into rat liver hepatocytes and endothelium. An *in vivo* autoradiographic study. *J. Cell Biol.* 80: 427-443.

21. Berman, J., Tachibana, C.Y. and Tye, B.-K. (1986). Identification of a telomere-binding activity from yeast. *Proc. Natl. Acad. Sci. USA* 83: 3713-3717.

22. Billet, F.S., Hamilton, L. and Newth, D.R. (1964). A failure to transform metazoan somatic cells by DNA. *Heredity* 19: 259-269.

23. Bishop, J.M. (1978). Retroviruses. *Ann. Rev. Biochem.* 47: 35-88.

24. Bishop, J.M. (1987). The molecular genetics of cancer. *Science* 235: 305-311.

25. Blundell, T.L., Dodson, G.G., Dodson, E., Hodgkin, D.C. and Vijayan, M. (1971). X-ray analysis and the structure of insulin. *Recent Prog. Hormone Res.* 27: 1-40.

26. Blundell, T.L., Dodson, G.G., Hodgkin, D.C. and Mercola, D. (1972). Insulin: the structure in the crystal and its reflection in chemistry and biology. *Adv. Protein Chem.* 26: 279-402.

27. Blundell, T. and Wood, S. (1982). The conformation, flexibility and dynamics of polypeptide hormones. *Ann. Rev. Biochem.* 51: 123-154.

28. Bolivar, F. and Backman, K. (1979). Plasmids of *Escherichia coli* as cloning vectors. *Methods in Enzymology* 68: 245-267.

29. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L. and Boyer, H.W. (1977). Construction and characterisation of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2: 95-113.

30. Boogaard, C. and Dixon, G.H. (1983). Red cell ghost-mediated microinjection of RNA into HeLa cells. II. Cellular translation of protamine mRNA; post-translational modifications and nuclear binding of newly synthesized protamine. *Exp. Cell Res.* 143: 191-205.

31. Borgmeyer, U., Nowock, J. and Sippel, A.E. (1984). The TGGCA-binding protein: a eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded linear DNA. *Nuc. Acids Res.* 12: 4295-4311.

32. Bosselman, R.A., Hsu, R.-Y., Boggs, T., Hu, S., Bruszewski, J., Ou, S., Kozar, L., Martin, F., Green, C., Jacobsen, F., Nicolson, M., Schultz, J.A., Semon, K.M., Rishell, W. and Stewart, R.G. (1989). Germline transmission of exogenous genes in the chicken. *Science* 243: 533-535.

33. Bowen, B., Steinberg, J., Laemmli, U.K. and Weintraub, H. (1980). The detection of DNA-binding proteins by protein blotting. *Nuc. Acids Res.* 8: 1-20.

34. Boynton, J.E., Gillham, N.W., Harris, E.H., Hosler, J.P., Johnson, A.M., Jones, A.R., Randolph-Anderson, B.L., Robertson, D., Klein, T.M., Shark, K.B. and Sanford, J.C. (1988). Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. *Science* 240: 1534-1538.

35. Brown, J.R. (1977). Structure and evolution of serum albumin. In *Albumin: Structure, Biosynthesis, Function*, ed. Peters, T. and Sjoholm, I. Colloquium B9, Vol. 50, pp.1-10, FEBS eleventh meeting, Copenhagen.

36. Brunfeldt, K., Hansen, B.A. and Jorgensen, K.R. (1968). The immunological reactivity and biological activity of iodinated insulin. *Acta Endocrinol.* 57: 307-329.

37. Cahill, G.F.Jr. (1971). Physiology of insulin in man. *Diabetes* 20: 785-799.

38. Capecchi, M.R. (1980). High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell* 22: 479-488.

39. Capecchi, M.R. (1989). Altering the genome by homologous recombination. *Science* 244: 1288-1292.

40. Carraway, K.L. and Koshland, D.E.Jr. (1972). Carbodiimide modification of proteins. *Methods in Enzymology* 25B: 616-623.

41. Celis, J.E., Kaltoft, K. and Bravo, R. (1980). Microinjection of somatic cells. In *Introduction of Macromolecules into Viable Mammalian Cells*, ed. Baserga, R., Croce, C. and Rovera, G. The Wistar Symposium Series Vol.1. pp.99-123. Alan R. Liss, Inc., New York.

42. Celis, J.E. (1984). Microinjection of somatic cells with micropipettes: comparison with other techniques. *Biochem. J.* 223: 281-291.

43. Cepko, C.L., Roberts, B.E. and Mulligan, R.C. (1984). Construction and applications of a highly transmissible murine retrovirus shuttle vector. *Cell* 37: 1053-1062.

44. Chambers, R. and Fell, H.B. (1931). Micro-operations on cells in tissue cultures. *Proc. Roy. Soc. Lond. B* 109: 380-403.

45. Cheng, S.-Y., Merlino, G.T. and Pastan, I.H. (1983). A versatile method for the coupling of proteins to DNA: synthesis of α_2-macroglobulin-DNA conjugates. *Nuc. Acids Res.* 11: 659-669.

46. Ciechanover, A., Schwartz, A.L. and Lodish, H.F. (1983). The asialoglycoprotein receptor internalises and recycles independently of the transferrin and insulin receptors. *Cell* 32: 267-275.

47. Cline, M.J., Stang, H., Mercola, K., Morse, L., Ruprecht, R., Browne, J. and Salser, W. (1980). Gene transfer in intact animals. *Nature* 284: 422-425.

48. Cohen, S.N., Chang, A.C.Y. and Hsu, L. (1972). Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA* 69: 2110-2114.

49. Colbere-Garapin, F., Horodniceanu, F., Kourilsky, P. and Garapin, A.-C. (1981). A new dominant hybrid selective marker for higher eukaryotic cells. *J. Mol. Biol.* 150: 1-14.

50. Cooper, D.J., Daniels, P.J.L., Yudis, M.D., Marigliano, H.M., Guthrie, R.D. and Bukhari, S.T.K. (1971). The gentamicin antibiotics. 3. The gross structures of gentamicin C components. *J. Chem. Soc. (C)* 1971: 3126-3129.

51. Costantini, F., Chada, K. and Magram, J. (1986). Correction of murine ß-thalassemia by gene transfer into the germ line. *Science* 233: 1192-1194.

52. Cox, M.M. and Lehman, I.R. (1987). Enzymes of general recombination. *Ann. Rev. Biochem.* 56: 229-262.

53. Cuatrecasas, P. (1969). Interaction of insulin with the cell membrane: the primary action of insulin. *Proc. Natl. Acad. Sci. USA* 63: 450-457.

54. Cuatrecasas, P. (1971). Insulin-receptor interactions in adipose tissue cells: direct measurement and properties. *Proc. Natl. Acad. Sci. USA* 68: 1264-1268.

55. Cuatrecasas, P., Hollenberg, M.D., Chang, K.-J. and Bennett, V. (1975). Hormone receptor complexes and their modulation of membrane function. *Recent Prog. Hormone Res.* 31: 37-94.

56. Cudd, A. and Nicolau, C. (1984). Entrapment of recombinant DNA in liposomes and its transfer and expression in eukaryotic cells. In *Liposome Technology,* ed. Gregoriadis, G. Vol.II, pp. 207-221. CRC Press Inc.

57. Cutfield, J.F., Cutfield, S.M., Dodson, E.J., Dodson, G.G., Emdin, S.F. and Reynolds, C.D. (1979). structure and biological activity of hagfish insulin. *J. Mol. Biol.* 132: 85-100.

58. Daniels, P.J.L., Yehaskel, A.S. and Morton, J.B. (1973). The structure of antibiotic G418. Thirteenth interscience conference on antimicrobial agents and chemotherapy, Washington. Abs. 137.

59. Darlington, G.J. (1987). Liver cell lines. *Methods in Enzymology* 151: 19-38.

60. Davies, J., Berg, D., Jorgensen, R., Fiandt, M. Huang, T.-S.R., Courvalin, P. and Schloff, J. (1977). R-factors: their properties and possible control. In *Topics in Infectious Diseases*, ed, Drews, J. and Hogenauer, G. Vol.2, pp. 101-114, Springer, New York.

61. Davies, J. and Smith, D.I. (1978). Plasmid-determined resistance to antimicrobial agents. *Ann. Rev. Microbiol.* 32: 469-518.

62. Davis, L.E., Roth, S.A. and Anderson, B. (1984). Antisera specificities to 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide adducts of proteins. *Immunology* 53: 435-441.

63. Deamer, D. and Bangham, A.D. (1976). Large volume liposomes by an ether vaporization method. *Biochim. Biophys. Acta* 443: 629-634.

64. De Block, M., Botterman, J., Vandewiele, M., Dockx, J., Thoen, C., Gossele, V., Raomovva, N., Thompson, C., Van Montagu, M. and Leemans, J. (1987). Engineering herbicide resistance in plants by expression of a detoxifying enzyme. *EMBO J.* 6: 2513-2518.

65. De La Pena, A., Lorz, H. and Schell, J. (1986). Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nature* 325: 274-276.

66. Derewenda, U., Derewenda, Z., Dodson, E.J., Dodson, G.G., Reynolds, C.D., Smith, G.D., Sparks, C. and Swenson, D. (1989). Phenol stabilises more helix in a new symmetrical zinc insulin hexamer. *Nature* 338: 594-596.

67. Diakumacos, E.G. (1980). Introduction of macromolecules into viable mammalian cells by precise physical microinjection. In *Introduction of Macromolecules into Viable Mammalian Cells*, ed, Baserga, R., Croce, C. and Rovera, G. The Wistar Symposium Series Vol.1. pp. 85-98. Alan R. Liss, Inc., New York.

68. Diakumacos, E.G., Holland, S. and Pecora, P. (1970). A microsurgical methodology for human cells *in vitro*: evolution and applications. *Proc. Natl. Acad. Sci. USA* 65: 911- 918.

69. Dick, J.E., Magli, M.C., Huszar, D., Phillips, R.A. and Bernstein, A. (1985). Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hemopoietic system of W/W^{v} mice. *Cell* 42: 71-79.

70. Domingo, D.L. and Trowbridge, I.S. (1985). Transferrin receptor as a target for antibody-drug conjugates. *Methods in Ebzymology* 112: 238-247.

71. Dubnau, D., Goldthwaite, D., Smith, I. and Marmur, J. (1967). Genetic mapping in *Bacillus subtilis*. *J. Mol. Biol.* 27: 163-185.

72. Duesberg, P.H. (1987). Latent cellular oncogenes: the paradox dissolves. *J. Cell Sci. Suppl.* 7: 169-187.

73. Dzierzak, E.A., Papayannopoulos, T. and Mulligan, R.C. (1988). Lineage-specific expression of a human ß-globin gene in murine bone marrow transplant recipients reconstituted with retrovirus transduced stem cells. *Nature* 331: 35-41.

74. Ebina, Y., Ellis, L., Jarnagin, K., Ederly, M., Graf, L., Clauser, E., Ou, J.-H., Masiarz, F., Kan, Y.W., Goldfine, I.D., Roth, R.A. and Rutter, W.J. (1985). The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling. *Cell* 40: 747-758.

75. Egelman, E.H. and Stasiak, A. (1986). Structure of helical recA-DNA complexes. Complexes formed in the presence of ATP-Gamma-S or ATP. *J. Mol. Biol.* 191: 677-697.

76. Elshourbagy, N.A., Walker, D.W., Paik, Y.-K., Boguski, M.S., Freeman, M., Gordan, J.I. and Taylor, J.M. (1987). Structure and expression of the human apolipoprotein A-IV gene. *J. Biol. Chem.* 262: 7973-7981.

77. Enquist, L.W., Van de Woude, G.F., Wagner, M., Smiley, J.R. and Summers, W.C. (1979). Construction and characterization of a recombinant plasmid encoding the gene for the thymidine kinase of *Herpes simplex* Type 1 virus. *Gene* 7: 335-342.

78. Erlanger, B.F. (1980). The preparation of antigenic hapten-carrier conjugates: a survey. *Methods in Enzymology* 70: 85-104.

79. Farber, F.E., Melnick, J.L. and Butel, J.S. (1975). Optimal conditions for uptake of exogenous DNA by Chinese hamster lung cells deficient in hypoxanthine-guanine phophoribosyltransferase. *Biochim. Biophys. Acta* 390: 298-311.

80. Faulstich, H. and Fiume, L. (1985). Protein conjugates of fungal toxins. *Methods in Enzymology* 112A: 225-237.

81. Fechheimer, M., Boylan, J.F., Parker, S., Sisken,. J.E., Patel, G.L. and Zimmer, S.G. (1987). Transfection of mammalian cells with plasmid DNA by scrape-loading and sonication loading. *Proc. Natl. Acad. Sci USA* 84: 8463-8467.

82. Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* 132: 6-13.

83. Feinberg, A.P. and Vogelstein, B. (1984). ADDENDUM: A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* 137: 266-267.

84. Flyer, D.C., Burakoff, S.J. and Faller, D.V. (1983). Cytotoxic T lymphocyte recognition of transfected cells expressing a cloned retroviral gene. *Nature* 305: 815-818.

85. Fraley, R.T., Fornari, C.S. and Kaplan, S. (1979). Entrapment of a bacterial plasmid in phospholipid vesicles: potential for gene transfer. *Proc. Natl. Acad. Sci. USA* 76: 3348-3352.

86. Fraley, R., Straubinger, R.M., Rule, G., Springer, E.L. and Papahadjopoulos, D. (1981). Liposome-madiated delivery of deoxyribonucleic acid to cells: enhanced efficiency of delivery related to lipid composition and incubation conditions. *Biochemistry* 20: 6978-6987.

87. Freychet, P., Roth, J. and Neville, D.M.Jr. (1971). Monoiodoinsulin: demonstration of its biological activity and binding to fat cells and liver membranes. *Biochem. Biophys. Res. Comm.* 43: 400-408.

88. Fried, M.G. and Crothers, D.M. (1981). Equilibrium kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nuc. Acids Res.* 9: 6505-6525.

89. Friedmann, T. (1989). Progress towards human gene therapy. *Science* 244: 1275-1281.

90. Fromm, M.E., Taylor, L.P. and Walbot, V. (1986). Stable gene transformation of maize after gene transfer by electroporation. *Nature* 319: 791-793.

91. Furusawa, M., Nishimura, T., Yamaizumi, M. and Okada, Y. (1974). Injection of foreign substances into single cells by cell fusion. *Nature* 249: 449-450.

92. Gammeltoft, S. (1984). Insulin receptors: binding kinetics and structure-function relationship of insulin. *Physiol. Revs.* 64: 1321-1378.

93. Gammeltoft, S., Kristensen, L.O. and Sestoft, L. (1978). Insulin receptors in isolated rat hepatocytes. Reassessment of binding properties and observations on the inactivation of insulin at 37°C. *J. Biol. Chem.* 253: 8406-8413.

94. Garrat, C.J., Harrison, D.M. and Wicks, M. (1972). The effect of iodination of particular tyrosine residues on the hormonal activity of insulin. *Biochem. J.* 126: 123-131.

95. Gartler, S.M. (1959). Cellular uptake of deoxyribonucleic acid by human tissue culture cells. *Nature* 184: 1505-1506.

96. Gartler, S.M. (1960). *Biochem. Biophys. Res. Comm.* 3: 127-

97. Gasser, C.S. and Fraley, R.T. (1989). Genetically engineering plants for crop improvement. *Science* 244: 1293-1299.

98. Gaudray, P., Tyndall, C., Kamen, R. and Cuzin, F. (1981). The high affinity binding site on polyoma virus DNA for the viral large-T protein. *Nuc. Acids Res.* 9: 5697-5710.

99. Gavin, J.R.III., Roth, J., Neville, D.M.Jr., De Meyts, P. and Buell, D.N. (1974). Insulin-dependent regulation of insulin receptor concentrations. A direct demonstration in cell culture. *Proc. Natl. Acad. Sci. USA* 71: 84-88.

100. Gerard, R.D. and Gluzman, Y. (1985). New host cell system for regulated simian virus 40 DNA replication. *Mol. Cell Biol.* 5: 3231-3240.

101. Gething, M.J. and Sambrook, J. (1981). Cell surface expression of influenza haemagglutinin from a cloned DNA copy of the RNA gene. *Nature* 293: 620-625.

102. Ghosh, P.K., Lebowitz, P., Frisque, R.J. and Gluzman, Y. (1981). Identification of a promoter component involved in positioning the 5′ termini of simian virus 40 early mRNAs. *Proc. Natl. Acad. Sci. USA* 78: 100-104.

103. Gitman, A.G., Graessmann, A. and Loyter, A. (1985). Targeting of loaded Sendai virus envelopes by covalently attached insulin molecules to virus receptor-depleted cells: fusion-mediated microinjection of ricin A and simian virus 40 DNA. *Proc. Natl. Acad. Sci. USA* 82: 7309-7313.

104. Glover, D.M. (ed.). (1985). *DNA Cloning.* Vol.I. IRL Press, Oxford.

105. Goldfine, I.D., Purello, F., Vigneri, R. and Clawson, G.A. (1985). Direct regulation of nuclear functions by insulin: relationship to mRNA metabolism. In *Molecular Basis of Insulin Action*, ed. Czech, M.P. pp. 329-345. Plenum Press, New York.

106. Goldstein, J.L., Anderson, R.G. and Brown, M.S. (1979). Coated pits, coated vesicles and receptor-mediated endocytosis. *Nature* 279: 679-685.

107. Goodfriend, T.L., Levine, L. and Fasman, G.D. (1964). Antibodies to bradykinin and angiotensin: a use of carbodiimides in immunology. *Science* 144: 1344-1346.

108. Gordon, J.W. and Ruddle, F.H. (1983). Gene transfer into mouse embryos: production of transgenic mice by pronuclear injection. *Methods in Enzymology* 101: 411-433.

109. Gorman, C. (1985). High efficiency gene transfer into mammalian cells. In *DNA Cloning*, ed. Glover, D.M. Vol.II. pp. 143-190. IRL Press, Oxford.

110. Gorman, C., Howard, B. and Reeves, R. (1983). Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nuc. Acids Res.* 11: 7631-7648.

111. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell Biol.* 2: 1044-1051.

112. Graessmann, A. (1970). Mikrochirurgische zellkern-transplantation bei saugetierzellen. *Exp. Cell Res.* 60: 373-382.

113. Graessmann, M. and Graessmann, A. (1976). "Early" simian-virus-40-specific RNA contains information for tumor antigen formation and chromatin replication. *Proc. Natl. Acad. Sci. USA* 73: 366-370.

114. Graessmann, M. and Graessmann, A. (1983). Microinjection of tissue culture cells. *Methods in Enzymology* 101: 482-492.

115. Graham, F.L., Bacchetti, S., McKinnon, R., Stanners, C., Cordell, B. and Goodman, H.M. (1980). Transformation of mammalian cells with DNA using the calcium technique. In *Introduction of Macromolecules into Mammalian Cells*, ed. Baserga, R., Croce, C. and Rovera, G. The Wistar Symposium Series Vol.1. pp. 3-25. Alan R. Liss, Inc., New York.

116. Graham, F.L. and Van der Eb, A.J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52: 456-461.

117. Gregoriadis, G. (1977). Targeting of drugs. *Nature* 265: 407-411.

118. Gregoriadis, G. (ed.). (1984). *Liposome Technology*. Vols. II and III. CRC Press, Florida.

119. Gregoriadis, G. and Buckland, R.A. (1973). Enzyme-containing liposomes alleviate a model for storage disease. *Nature* 244: 170-172.

120. Gregoriadis, G. and Meehan, A. (1981). Interaction of antibody-bearing small unilamellar liposomes with antigen-coated cells. *Biochem. J.* 200: 211-216.

121. Griffith, F. (1928). The significance of pneumococcal types. *J. Hygiene* 27: 113-159.

122. Gruss, P., Dhar, R. and Khoury, G. (1981). Simian virus 40 tandem repeated sequences as an element of the early promoter. *Proc. Natl. Acad. Sci. USA* 78: 943-947.

123. Gruss, P. and Khoury, G. (1981). Expression of simian virus 40-rat proinsulin recombinants in monkey kidney cells: use of preproinsulin RNA processing signals. *Proc. Natl. Acad. Sci. USA* 78: 133-137.

124. Guroff, G. (Ed.). (1987). *Oncogenes, Genes and Growth Factors*. J. Wiley and Sons, New York.

125. Hall, C.V., Jacob, P.E., Ringold, G.M. and Lee, F. (1983). Expression and regulation of *Escherichia coli lacZ* gene fusions in mammalian cells. *J. Mol. Appl. Genet.* 2: 101-109.

126. Hammer, R.E., Pursel, V.G., Rexroad, C.E.Jr., Wall, R.J., Bolt, D.J., Ebert, K.M., Palmiter, R.D. and Brinster, R.L. (1985). Production of transgenic rabbits, sheep and pigs by microinjection. *Nature* 315: 680-683.

127. Hanahan, D. (1985). Techniques for the transformation of *E. coli*. In *DNA Cloning*, ed. Glover, D.M. Vol.I. pp. 109-135. IRL Press, Oxford.

128. Haralambidis, J., Duncan, L., Angus, K., Wade, J. and Tregear, G. (1989). Oligonucleotide-peptide hybrids - a possible means of targeting oligonucleotides. Conference on oligonucleotides as antisense inhibitors of gene expression: therapeutic implications, Rockville, Md., USA.

129. Haring, H.U., White, M.F., Kahn, C.R., Kasuga, M., Lauris, V., Fleischmann, R., Murray, M. and Pawelek, J. (1984). Abnormality of insulin binding and receptor phosphorylation in an insulin-resistant melanoma cell line. *J. Cell Biol.* 99: 900-908.

130. Harley, E.H., Oliver, C.L., Rhodes-Harrison, L., Mew, R.T., Lecatsas, G. and Naude, W.duT. (1982). Cloning and characterisation of BKV(MM) DNA and its use for detection of BKV DNA in human urine. *S. A. J. Sci.* 78: 112-116.

131. Hasezawa, S., Nagata, T. and Syono, K. (1981). Transformation of *Vinca* protoplasts mediated by *Agrobacterium* spheroplasts. *Mol. Gen. Genet.* 182: 206-210.

132. Hecht, M.H., Nelson, H.C.M. and Sauer, R.T. (1983). Mutations in λ repressor's amino-terminal domain: implications for protein stability and DNA binding. *Proc. Natl. Acad. Sci. USA* 80: 2676-2680.

133. Heidenreich, K.A. and Olefsky, J.M. (1985). The metabolism of insulin receptors: internalization, degradation and recycling. In *The Molecular Basis of Insulin Action*, ed. Czech, M.P. pp. 45-65. Plenum Press, New York.

134. Hennighausen, L. and Lubon, H. (1987). Interaction of proteins with DNA *in vitro*. *Methods in Enzymology* 152: 721-735.

135. Hirt, B. (1967). Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* 26: 365-369.

136. Hjelle, B., Liu, E. and Bishop, J.M. (1988). Oncogene v-*src* transforms and establishes embryonic rodent fibroblasts but not diploid human fibroblasts. *Proc. Natl. Acad. Sci. USA* 85: 4355-4359.

137. Hock, R.A. and Miller, A.D. (1986). Retrovirus-mediated transfer and expression of drug resistance genes in human haematopoietic progenitor cells. *Nature* 320: 275-277.

138. Holmes, D.S. and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Analyt. Biochem.* 114: 193-197.

139. Hooper, I.R. (1982). The naturally occurring aminoglycoside antibiotics. In *Aminoglycoside Antibiotics*, ed. Umezawa, H. and Hooper, I.R. pp. 1-35. Springer-Verlag, Berlin.

140. Hopwood, D.A. and Wright, H.M. (1972). Transformation in *Thermoactinomyces vulgaris*. *J. Gen. Microbiol.* 71: 383-398.

141. Hosaka, Y. and Shimizu, Y.K. (1972). Artificial assembly of envelope particles of HVJ (Sendai virus). I. Assembly of hemolytic and fusion factors from envelope solubilized by Nonidet P40. *Virology* 49: 627-639.

142. Hotchkiss, R.D. and Gabor, M. (1970). Bacterial transformation, with special reference to recombination process. *Ann. Rev. Genet.* 4: 193-224.

143. Huckett, B., Gordhan, H., Hawtrey, R., Moodley, N., Ariatti, M. and Hawtrey, A. (1986). Binding of DNA to albumin and transferrin modified by treatment with water-soluble carbodiimides. *Biochem. Pharmacol.* 35: 1249-1257.

144. Hunter, W.H. and Greenwood, F.C. (1962). Preparation of iodine-131 labelled human growth hormone of high specific activity. *Nature* 194: 495-496.

145. Iino, T., Furusawa, M., Furusawa, I. and Obinata, M. (1983). Transformation of L cells with virus thymidine kinase genes introduced by red cell-mediated microinjection. *Exp. Cell Res.* 148: 475-480.

146. Isa, K. and Shima, A. (1987). Transfection and stable expression of a dominant selective marker *Ecogpt* in a cultured cell line of the fish *Carassius auratus*. *J. Cell Sci.* 88: 219-224.

147. Ishimi, Y., Claude, A. Bullock, P. and Hurwitz, J. (1988). Complete enzymatic synthesis of DNA containing the SV40 origin of replication. *J. Biol. Chem.* 263: 19723-19733.

148. Ivanov, V.O., Preobrazhensky, S.N., Tsibulsky, V.P., Babaev, V.R., Repin, V.S. and Smirnov, V.N. (1985). Liposome uptake by cultured macrophages mediated by modified low-density lipoproteins. *Biochim. Biophys. Acta* 846: 76-84.

149. Izzo, J.L., Roncone, A., Izzo, M.J. and Bale, W.F. (1964). Relationship between the degree of iodination of insulin and its biological, electrophoretic and immunochemical properties. *J. Biol. Chem.* 239: 3749-3754.

150. Jackson, D.A. and Cook, P.R. (1986). Replication occurs at a nucleoskeleton. *EMBO J.* 5: 1403-1410.

151. Jackson, P.D. and Felsenfeld, G. (1987). *In vivo* footprinting of specific protein-DNA interactions. *Methods in Enzymology* 152: 735-755.

152. Jacobs, S. and Cuatrecasas, P. (1986). Phosphorylation of receptors for insulin and insulin-like growth factor I: effects of hormones and phorbol esters. *J. Biol. Chem.* 261: 934-939.

153. Jacobs, S., Hazim, E. and Cuatrecasas, P. (1980). The subunit structure of rat liver insulin receptor. Antibodies directed against the insulin-binding subunit. *J. Biol. Chem.* 255: 6937-6940.

154. Jaenisch, R. (1988). Transgenic animals. *Science* 240: 1468-1474.

155. Jakob, H. and Nicolas, J.-F. (1987). Mouse teratocarcinoma cells. *Methods in Enzymology.* 151: 66-81.

156. Jarett, L. and Smith, R.M. (1974). Electron microscopic demonstration of insulin receptors on adipocyte plasma membrane utilizing a ferritin-insulin conjugate. *J. Biol. Chem.* 249: 7024-7031.

157. Jefferson, R.A. Kavanagh, T.A. and Bevan, M.W. (1987). GUS fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* 6: 3901-3907.

158. Jensen, D.E. and Von Hippel, P.H. (1976). DNA "melting" proteins. I. Effects of bovine pancreatic ribonuclease binding on the conformation and stability of DNA. *J. Biol. Chem.* 251: 7198-7214.

159. Jensen, D.E., Kelly, R.C. and Von Hippel, P.H. (1976). DNA "melting" proteins. II. Effects of bacteriophage T4 gene 32-protein binding on the conformation and stability of nucleic acid structures. *J. Biol. Chem.* 251: 7215-7228.

160. Ji, T.H. (1983). Bifunctional reagents. *Methods in Enzymology* 91: 580-609.

161. Jimenez, A. and Davies, J. (1980). Expression of a transposable antibiotic resistance element in *Saccharomyces*. *Nature* 287: 869-871.

162. Johansen, H., Reff, M., Schumperli, D. and Rosenberg, M. (1984). Studying eukaryotic transcription and translation by gene fusion. In *Gene Expression. The Translational Step and its Control*, ed. Clark, B.F.C. and Petersen, H.U. pp. 413-426. Munksgaard, Copenhagen.

163. Johnson, A.D., Pabo, C.O. and Sauer, R.T. (1980). Bacteriophage λ repressor and cro protein: interactions with operator DNA. *Methods in Enzymology* 65: 839-856.

164. Johnston, S.A., Anziano, P.Q., Shark, K., Sanford, J.C. and Butow, R.A. (1988). Mitochondrial transformation in yeast by bombardment with microprojectiles. *Science* 240: 1538-1541.

165. Jones, J., Dunsmuir, P. and Bedbrook, J. (1985). High level expression of introduced chimaeric genes in regenerated transformed plants. *EMBO J.* 4: 2411-2418 (1985).

166. Jorgensen, R.A., Rothstein, S.J. and Reznikoff, W.S. (1979). A restriction enzyme cleavage map of Tn5 and location of a region encoding neomycin resistance. *Mol. Gen. Genet.* 177: 65-72.

167. Kaneda, Y., Iwai, K. and Uchida, T. (1989). Increased expression of DNA cointroduced with nuclear protein in adult rat liver. *Science* 243: 375-378.

168. Karpel, R.L., Merkler, D.J., Flowers, B.K. and Delahunty, M.D. (1981). Involvement of basic amino acids in the activity of a nucleic acid helix-destabilizing protein. *Biochim. Biophys. Acta* 654: 42-51.

169. Karube, I., Tamiya, E. and Matsuoka, H. (1985). Transformation of *Saccharomyces cerevisiae* spheroplasts by high electric pulse. *FEBS Lett.* 182: 90-94.

170. Kato, S., Anderson, R.A. and Camerini-Otero, R.D. (1986). Foreign DNA introduced by calcium phosphate is integrated into repetitive DNA elements of the mouse L cell genome. *Mol. Cell Biol.* 6: 1787-1795.

171. Katz, L., Kingsbury, D. and Helinski, D.R. (1973). Stimulation by cyclic adenosine monophosphate of plasmid deoxyribonucleic acid replication and catabolite repression of the plasmid deoxyribonucleic acid-protein relaxation complex. *J. Bact.* 114: 577-591.

172. Kawai, S. and Nishizawa, M. (1984). New procedure for DNA transfection with polycation and dimethyl sulfoxide. *Mol. Cell Biol.* 4: 1172-1174

173. Keefer, L.M., Piron, M.-A. and De Meyts, P. (1981). Human insulin prepared by recombinant DNA techniques and native human insulin interact identically with insulin receptors. *Proc. Natl. Acad. Sci. USA* 78: 1391-1395.

174. Keen, J.H. (1985). The structure of clathrin-coated membranes: assembly and disassembly. In *Endocytosis*, ed. Pastan, I. and Willingham, M.C. pp.85-130. Plenum Press, New York.

175. Kerbel, R.S., Waghorne, C., Man, M.S., Elliot, B. and Breitman, M.L. (1987). Alteration of the tumorigenic and metastatic properties of neoplastic cells is associated with the process of calcium-phosphate-mediated DNA transfection. *Proc. Natl. Acad. Sci. USA* 84: 1263-1267.

176. Khorana, H.G. (1953). The chemistry of carbodiimides. *Chem. Revs.* 53: 145-166.

177. Klein, T.M., Wolf, E.D., Wu, R. and Sanford, J.C. (1987). High velocity microprojectiles for delivering nucleic acids into living cells. *Nature* 327: 70-73.

178. Knowles, B.B., Howe, C.C. and Aden, D.P. (1980). Human hepatocellular carcinoma cell lines secrete the major plasma proteins and Hepatitis B surface antigen. *Science* 209: 497-499.

179. Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y. and Inouye, K. (1982). Supernormal insulin: [D-Phe^{B24}]-insulin with increased affinity for insulin receptors. *Biochem. Biophys. Res. Comm.* 107: 329-336.

180. Kono, T. (1975). Determination of the binding characteristics of insulin to fat cells. *Methods in Enzymology* 37B: 193-198.

181. Kopczynski, M.G. and Babior, B.M. (1984). The mechanism of action of ethanolamine ammonia-lyase, an adenosylcobalamin-dependent enzyme. Evidence for a carboxyl group at the active site. *J. Biol. Chem.* 259: 7652-7654.

182. Kraus, L.M. (1961). Formation of different haemoglobins in tissue culture of human bone marrow treated with human deoxyribonucleic acid. *Nature* 192: 1055-1057.

183. Kuhn, S., Vorgias, C.E. and Traub, P. (1987). Interaction *in vitro* of non-epithelial intermediate filament proteins with supercoiled plasmid DNA. *J. Cell Sci.* 87: 543-554.

184. Kurtz, D.I. (1974). Fidelity of protein synthesis with chicken embryo mitochondrial and cytoplasmic ribosomes. *Biochemistry* 13: 572-577.

185. Kushner, Kh.F. (1957). Effect of metabolic factors on heredity in animals. *Agrobiologia* 1: 19-35.

186. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685.

187. Laimins, L.A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982). Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. *Proc. Natl. Acad. Sci. USA* 79: 6453-6457.

188. Land, H., Parada, J.T. and Weinberg, R.A. (1983). Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature* 304: 596-602.

189. Langley, J.N. (1905). On the reaction of cells and of nerve endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and curare. *J. Physiol. (London)* 33: 374-413.

190. Laval, F., Malaise, E. and Laval, J. (1970). Heterologous DNA nuclear uptake by mouse fibroblasts *in vitro* and its early fate. *Exp. Cell Res.* 63: 69-77.

191. Ledoux, L. (1965). Uptake of DNA by living cells. *Progress Nuc. Acid Res. and Mol. Biol.* 4: 231-267.

192. Ledoux, L. and Huart, R. (1968). Integration and replication of DNA of *M lysodeikticus* in DNA of germinating barley. *Nature* 1256-1259.

193. Ledoux, L and Huart, R. (1969). Fate of exogenous bacterial deoxyribonucleic acids in barley seedlings. *J. Mol. Biol.* 43: 243-262.

194. Lee, C.K. and Knipe, D.M. (1985). An immunoassay for the study of DNA-binding activities of Herpes Simplex Virus protein ICP8. *J. Virol.* 54: 731-738.

195. Leroy, P., Benoit, J., Vendrely, R. and Vendrely, C. (1964). *Compt. Rend. Acad. Sci.* 258: 1905-

196. Levinger, L. and Varshavsky, A. (1982). Protein D1 preferentially binds A+T-rich DNA *in vitro* and is a component of *Drosophila melanogaster* nuleosomes containing A+T-rich satellite DNA. *Proc. Natl. Acad. Sci. USA* 79: 7152-7156.

197. Levitzki, A. (1985). Receptors. In *Endocytosis*, ed. Pastan, I. and Willingham, M.C. pp. 45-68. Plenum Press, New York.

198. Linde, S., Hansen, B. and Lernmark, A. (1983). Preparation of stable radioiodinated polypeptide hormones and proteins using polyacrylamide gel electrophoresis. *Methods in Enzymology* 92: 309-335.

199. Littlefield, J.W. (1964). Selection of hybrids from matings of fibroblasts*in vitro* and their presumed recombinants. *Science* 145: 709-710.

200. Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* 193: 265-275.

201. Loyter, A., Scangos, G.A. and Ruddle, F.H. (1982). Mechanisms of DNA uptake by mammalian cells: fate of exogenously added DNA monitored by the use of fluorescent dyes. *Proc. Natl. Acad. Sci. USA* 79: 422-426.

202. Loyter, A., Zakai, N. and Kulka, R.G. (1975). "Ultramicroinjection" of macromolecules or small particles into animal cells. A new technique based on virus-induced cell fusion. *J. Cell Biol.* 66: 292-304.

203. Lurquin, P.F. and Sheehy, R.E. (1982). Binding of large liposomes to plant protoplasts and delivery of encapsulated DNA. *Plant Sci. Lett.* 25: 133-146.

204. Luthman, H. and Magnusson, G. (1983). High efficiency polyoma DNA transfection of chloroquine treated cells. *Nuc. Acids Res.* 11: 1295-1308.

205. Mackett, M., Smith, G.L. and Moss, B. (1982). Vaccinia virus: a selectable eukaryotic cloning and expression vector. *Proc. Natl. Acad. Sci. USA* 79: 7415-7419.

206. Maehr, H. and Schaffner, C.P. (1967). The chemistry of the gentamicins. I. Characterization and gross structure of gentamicin A. *J. Am. Chem. Soc.* 89: 6787-6788.

207. Mandel, M. and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* 53: 159-162.

208. Mandelstam, J., McQuillen, K. and Dawes, I. (Ed.). (1982). *The Biochemistry of Bacterial Growth.* 3rd. Edition. Blackwell Scientific Publications.

209. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). *Molecular Cloning, a Laboratory Manual.* Cold Spring Harbour Laboratory.

210. Maniatis, T., Goodbourn, S. and Fischer, J.A. (1987). Regulation of inducible and tissue specific gene expression. *Science* 236: 1237-1245.

211. Margulis, L. (1974). Five kingdom classification and the origin and evolution of cells. *Evol. Biol.* 7: 45-78.

212. Mariman, E.C.M., Van Eekelen, C.A.G., Reinders, R.J., Berns, A.J.M. and Van Venrooij, W.J. (1982). Adenoviral heterogeneous nuclear RNA is associated with the host nuclear matrix during splicing. *J. Mol. Biol.* 154: 103-119.

213. Marsh, M. and Helenius, A. (1980). Adsorptive endocytosis of Semliki Forest Virus. *J. Mol. Biol.* 142: 439-454.

214. Marshall, S. (1985). Kinetics of insulin receptor internalization and recycling in adipocytes: shunting of receptors to a degradastive pathway by inhibitors of recycling. *J. Biol. Chem.* 260: 4136-4144.

215. Martin, G.R. and Evans, M.J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies *in vitro*. *Proc. Natl. Acad. Sci. USA* 72: 1441-1445.

216. Martinovitch, P.N., Kanazir, D.T., Knezevitch, Z.A. and Simitch, M.M. (1961). *Compt. Rend. Acad. Sci.* 253: 1854-

217. Martinovitch, P.N., Kanazir, D.T., Knezevitch, Z.A. and Simitch, M.M. (1962). *J. Embryol. Exptl. Morphol.* 10: 167-

218. Mason, W.S., Taylor, J.M. and Hull, R. (1987). Retroid virus genome replication. *Adv. Virus Res.* 32: 35-96.

219. Massague, J., Pilch, P.F. and Czech, M.P. (1980). Electrophoretic resolution of three major insulin receptor structures with unique subunit stoichiometries. *Proc. Natl. Acad. Sci. USA* 77: 7137-7141.

220. Matsuhashi, Y., Sawa, T., Takeuchi, T. and Umezawa, H. (1976). Purification of aminoglycoside 3′-phosphotransferase II. *J. Antibiot. (Tokyo)* 29: 204-207.

221. Matsuhashi, Y., Yagisawa, M., Kondo, S., Takeuchi, T. and Umezawa, H. (1975). Aminoglycoside 3′-phosphotransferases I and II in *Pseudomonas aeruginosa*. *J. Antibiot. (Tokyo)* 28: 442-447.

222. Maxfield, F.R., Schlessinger, J., Shecter, Y., Pastan, I. and Willingham, M.C. (1978). Collection of insulin, EGF and α_2-macroglobulin in the same patches on the surface of cultured fibroblasts and common internalization. *Cell* 14: 805-810.

223. McCutchan, J.H. and Pagano, J.S. (1968). Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. *J. Natl. Cancer Inst.* 41: 351-357.

224. McKay, D.B. and Steitz, T.A. (1981). Structure of catabolite gene activator protein at 2.9Å resolution suggests binding to left-handed B-DNA. *Nature* 290: 744-749.

225. McLachlan, A.D. and Walker, J.E. (1977). Evolution of serum albumin. *J. Mol. Biol.* 112: 543-558.

226. McNeil, P.L., Murphy, R.F., Lanni, F. and Taylor, D.L. (1984). A method for incorporating macromolecules into adherent cells. *J. Cell Biol.* 98: 1556-1564.

227. McNeil, P.L. and Warder, E. (1987). Glass beads load macromolecules into living cells. *J. Cell Sci.* 88: 669-678.

228. Melancon, P., Burgess, R.R. and Record, M.T.Jr. (1982). Nitrocellulose filter binding studies of the interactions of *Escherichia coli* RNA polymerase holoenzyme with deoxyribonucleic acid restriction fragments: evidence for multiple classes of nonpromoter interactions, some of which display promoter-like properties. *Biochemistry* 21: 4318-4331.

229. Melton, D.W. (1987). Strategies and mechanisms for the control of transcriptional initiation of mammalian protein-coding genes. *J. Cell Sci.* 88: 267-270.

230. Merril, C.R., Geier, M.R. and Petricciani, J.C. (1971). Bacterial virus gene expression in human cells. *Nature* 233: 398-400.

231. Mignotte, B., Barat, M. and Mounlou, J.-C. (1985). Characterization of a mitochondrial protein binding to single-stranded DNA. *Nuc. Acids Res.* 13: 1703-1716.

232. Miller, A.D., Eckner, R.J., Jolly, D.J., Friedmann, T. and Verma, I.M. (1984). Expression of a retrovirus encoding human HPRT in mice. *Science* 225: 630-632.

233. Monsan, P., Puzo, G. and Mazarguil, H. (1975). On the mechanism of the formation of glutaraldehyde-protein bonds. *Biochimie* 57: 1281-1292.

234. Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. and Chambon, P. (1981). The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimaeric recombinants. *Nuc. Acids Res.* 9: 6047-6068.

235. Morikawa, H., Iida, A. and Yamada, Y. (1988). Electric gene transfer into plant protoplasts and cells. In *Biotechnology in Agriculture*, ed. Mizrahi, A. Advances in Biotechnological Processes Vol.9. pp. 175-202. Alan R Liss, Inc., New York.

236. Mueller, C., Graessmann, A. and Graessman, M. (1980). Mapping of simian virus 40 tumor antigens and their functions using the microinjection technique. In *Introduction of Macromolecules into Viable Mammalian Cells*, ed. Baserga, R., Croce, C. and Rovera, G. The Wistar Symposium Series Vol.1. pp. 135-144. Alan R. Liss, Inc., New York.

237. Mukherjee, A.B., Orloff, S., Butler, J.deB., Triche, T., Lalley, P. and Schulman, J.D. (1978). Entrapment of metaphase chromosomes into phospholipid vesicles (lipochromosomes): carrier potential in gene transfer. *Proc. Natl. Acad. Sci. USA* 75: 1361-1365.

238. Mulligan, R.C., Howard, B.H. and Berg, P. (1979). Synthesis of rabbit ß-globin in cultured monkey kidney cells following infection with a SV40 ß-globin recombinant genome. *Nature* 277: 108-114.

239. Mulligan, R.C. and Berg, P. (1980). Expression of a bacterial gene in mammalian cells. *Science* 209: 1422-1427.

240. Nabel, E.G., Plautz, G., Boyce, F.M., Stanley, J.C. and Nabel, G.J. (1989). Recombinant gene expression *in vivo* within the endothelial cells of the arterial wall. *Science* 244: 1342-1344.

241. Neumann. E. and Rosenheck, K. (1972). Permeability changes induced by electrical impulses in vesicular membranes. *J. Membr. Biol.* 10: 279-290.

242. Newport, J.W. and Forbes, D.J. (1987). The nucleus: structure, function and dynamics. *Ann. Rev. Biochem.* 56: 535-565.

243. Nickol, J.M. and Felsenfeld, G. (1988). Bidirectional control of the chicken ß- and ε-globin genes by a shared enhancer. *Proc. Natl. Acad. Sci. USA* 85: 2548-2552.

244. Nicolau, C. and Sene, C. (1982a). Liposome-mediated DNA transfer in eukaryotic cells. Dependence of the transfer efficiency upon the type of liposomes used and the host cell cycle stage. *Biochim. Biophys. Acta* 721: 185-190.

245. Nicolau, C. and Sene, C. (1982b). Liposome-mediated gene transfer in eukaryotic cells: gene uptake and expression in the host cell. In *Targeting of Drugs*, ed. Gregoriadis, G., Senior, J. and Trouet, A. p.393. Plenum Press, New York.

246. Notani, N.K. & Setlow, J.K. (1974). Mechanisms of bacterial transformation and transfection. *Progress Nuc. Acid Res. and Mol. Biol.* 14: 39-100.

247. Oka, T. and Topper, Y.J. (1974). A soluble super-active form of insulin. *Proc. Natl. Acad. Sci. USA* 71: 1630-1633.

248. Okuda, T. and Ito, Y. (1982). Biosynthesis and mutasynthesis of aminoglycoside antibiotics. In *Aminoglycoside Antibiotics*, ed. Umezawa, H. and Hooper, I.R. pp. 111-203. Springer-Verlag, Berlin.

249. Pabo, C.O., Krovatin, W., Jeffrey, A. and Sauer, R.T. (1982). The N-terminal arms of λ repressor wrap around the operator DNA. *Nature* 298: 441-443.

250. Pabo, C.O. and Lewis, M. (1982). the operator binding domain of λ repressor: structure and DNA recognition. *Nature* 298: 443-447.

251. Pabo, C.O. and Sauer, R.T. (1984). Protein-DNA recognition. *Ann. Rev. Biochem.* 53: 293-321.

252. Palmer, T.D., Hock, R.A., Osborne, W.R.A. and Miller, A.D. (1987). Efficient retrovirus-mediated transfer and expression of a human adenosine deaminase gene in diploid skin fibroblasts from an adenosine deaminase-deficient human. *Proc. Natl. Acad. Sci. USA* 84: 1055-1059.

253. Palmiter, R.D., Brinster, R.L., Hammer, R.E., Trumbauer, M.E., Rosenfeld, M.G., Birnberg, N.C. and Evans, R.M. (1982). Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* 300: 611-615.

254. Papaphilippou, A.P., Department of Pharmacy, University of Durban-Westville, Private Bag X54001, Durban, 4000, R.S.A.

255. Parsons, T.F. and Pierce, J.G. (1979). Biologically active covalently cross-linked glycoprotein hormones and the effects of modification of the COOH-terminal region of their α-subunits. *J. Biol. Chem.* 254: 6010-6015.

256. Pastan, I. and Willingham, M.C. (1985). The pathway of endocytosis. In *Endocytosis*, ed. Pastan, I. and Willingham, M.C. pp. 1-44. Plenum Press, New York.

257. Pessin, J.E., Mottola, C., Yu, K.-T. and Czech, M.P. (1985). Subunit structure and regulation of the insulin-receptor complex. In *The Molecular Basis of Insulin Action*, ed. Czech, M.P. pp. 3-29. Plenum Press, New York.

258. Pederson, D.S., Thoma, F. and Simpson, R.T. (1986). Core particle, fiber, and transcriptionally active chromatin structure. *Ann. Rev. Cell Biol.* 2: 117-147.

259. Peeters, P.A.M., Claessens, C.A.M., Eling, W.M.C. and Crommelin, D.J.A. (1988). Immunospecific targeting of liposomes to erythrocytes. *Biochem. Pharmacol.* 37: 2215-2222.

260. Peters, K. and Richards, F.M. (1977). Chemical cross-linking: reagents and problems in studies of membrane structure. *Ann. Rev. Biochem.* 46: 523-551.

261. Peters, T.Jr. and Reed, R.G. (1977). Serum albumin: conformation and active sites. In *Albumin: Structure, Biosynthesis, Function*, ed. Peters, T. and Sjoholm, I. Colloqium B9, Vol.50, pp. 11-20, FEBS eleventh meeting, Copenhagen.

262. Pimentel, E. (1986). *Oncogenes.* CRC Press, Florida.

263. Pine, R., Levy, D.E., Reich, N. and Darnell, J.E.Jr. (1988). Transcriptional stimulation by $CaPO_4$-DNA precipitates. *Nuc. Acids Res.* 16: 1371-1378.

264. Poste, G. and Nicolson, G.L. (Ed.) (1978). *Membrane Fusion.* Cell Surface Reviews, Volume 5. Elsevier/North-Holland Biomedical Press, Amsterdam, The Netherlands.

265. Potter, H., Weir, L. and Leder, P. (1984). Enhancer-dependent expression of human K immunoglobulin genes introduced into mouse pre-B lymphoctes by electroporation. *Proc. Natl. Acad. Sci. USA* 81: 7161-7165.

266. Poznansky, M.J. and Bhardwaj, D. (1981). Antibody-mediated targeting of α-1,4-glucosidase-albumin polymers to rat hepatocytes. *Biochem. J.* 196: 89-93.

267. Poznansky, M.J., Shandling, M., Salkie, M.A., Elliot, J. and Lau, E. (1982). Advantages in the use of L-asparaginase-albumin polymer as an antitumor agent. *Cancer Res.* 42: 1020-1025.

268. Poznansky, M.J., Singh, R., Singh, B. and Fantus, G. (1984). Insulin: carrier potential for enzyme and drug therapy. *Science* 223: 1304-1306.

269. Proudfoot, N.J. and Brownlee, G.G. (1976). 3′ Non-coding region sequences in eukaryotic messenger RNA. *Nature* 263: 211-214.

270. Ptashne, M. (1967). Specific binding of the λ phage repressor to λ DNA. *Nature* 214: 232-234.

271. Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* 322: 697-701.

272. Pullen, R.A., Lindsay, D.G., Wood, S.P., Tickle, I.J., Blundell, T.L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., Glieman, J. and Gammeltoft, S. (1976). Receptor-binding region of insulin. *Nature* 259: 369-373.

273. Pursel, V.G., Pinkert, C.A., Miller, K.F., Bolt, D.J., Campbell, R.G., Palmiter, R.D., Brinster, R.L. and Hammer, R.E. (1989). Genetic engineering of livestock. *Science* 244: 1281-1288.

274. Radloff, R., Bauer, W. and Vinograd, J. (1967). A dye-buoyant density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA in HeLa cells. *Pric. Natl. Acad. Sci. USA* 57: 1514-1521.

275. Razin, S.V. (1987). DNA interactions with the nuclear matrix and spatial organisation of replication and transcription. *BioEssays* 6: 19-23.

276. Record, M.T.Jr., Lohman, T.M. and de Haseth, P. (1976). Ion effects on ligand-nucleic acid interactions. *J. Mol. Biol.* 107: 145-158.

277. Reddy, V.B., Thimmappaya, B., Dhar, R., Subramanian, K.N., Zain, B.S., Pan, J., Ghosh, P.K., Celma, M.L. and Weissman, S.M. (1978). The genome of simian virus 40. *Science* 200: 494-502.

278. Reichlin, M. (1980). Use of glutaraldehyde as a coupling agent for proteins and peptides. *Methods in Enzymology* 70A: 159-165.

279. Reiss, B., Sprengel, R., Will, H. and Schaller, H. (1984). A new sensitive method for qualitative and qhantitative assay of neomycin phosphotransferase in crude cell extracts. *Gene* 30: 211-218.

280. Riehm, J.P. and Scheraga, H.A. (1966). Structural studies of ribonuclease. XXI. The reaction between ribonuclease and a water-soluble carbodiimide. *Biochemistry* 5: 99-115.

281. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113: 237-251.

282. Riggs, A.D., Suzuki, H. and Bourgeois, S. (1970). *lac* Repressor-operator interaction. I. Equilibrium studies. *J. Mol. Biol.* 48: 67-83.

283. Robertson, M. (1986). Gene therapy: desperate appliances. *Nature* 320: 213-214.

284. Rogers, S.G., Horsch, R.B. and Fraley, R.T. (1986). Gene transfer in plants: production of transformed plants using Ti plasmid vectors. *Methods in Enzymology* 118: 627-640.

285. Rose, J.K. and Bergmann, J.E. (1982). Expression from cloned cDNA of cell surface secreted forms of the glycoprotein of vesicualr stomatitis virus in eucaryotic cells. *Cell* 30: 753-762.

286. Rosen, O.M. (1987). After insulin binds. *Science* 237: 1452-1458.

287. Rosenberg, M.B., Breakfield, X.O. and Hawrot, E. (1987). Targeting of liposomes to cells bearing nerve growth factor receptors mediated by biotinylated nerve growth factor. *J. Neurochem.* 48: 865-875.

288. Rosenthal, N. (1987). Identification of regulatory elements of cloned genes with functional assays. *Methods in Enzymology* 152: 704-720.

289. Roth, J. (1973). Peptide hormone binding to receptors: a review of direct studies *in vitro*. *Metabolism* 22: 1059-1073.

290. Rothstein, S.J., Jorgensen, R.A., Postle, K. and Reznikoff, W.S. (1980). The inverted repeats of Tn5 are functionally different. *Cell* 19: 795-805.

291. Rubenstein, J.L.R., Nicolas, J.-F. and Jacob, F. (1986). Introduction of genes into preimplantation mouse embryos by use of a defective recombinant retrovirus. *Proc. Natl. Acad. Sci. USA* 83: 366-368.

292. Ryle, A.P., Sanger, F., Smith, L.F. and Kitai, R. (1955). The disulphide bonds of insulin. *Biochem. J.* 60: 541-556.

293. Sancar, A. and Sancar, G.B. (1988). DNA repair enzymes. *Ann. Rev. Biochem.* 57: 29-67.

294. Sandri-Goldin, R.M., Goldin, A.L., Levine, M. and Glorioso, J.C. (1981). High-frequency transfer of cloned herpes simplex virus type 1 sequences to mammalian cells by protoplast fusion. *Mol. Cell Biol.* 1: 743-752.

295. Sandri-Goldin, R.M., Goldin, A.L., Levine, M. and Glorioso, J. (1983). High-efficiency transfer of DNA into eukaryotic cells by protoplast fusion. *Methods in Enzymology* 101: 402-411.

296. Sauer, R.T., Yocum, R.R., Doolittle, R.F., Lewis, M. and Pabo, C.O. (1982). Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. *Nature* 298: 447-451.

297. Schaefer-Ridder, M., Wang, Y. and Hofschneider, P.H. (1982). Liposomes as gene carriers: efficient transformation of mouse L cells by thymidine kinase gene. *Science* 215: 166-168.

298. Schaffner, W. (1980). Direct transfer of cloned genes from bacteria to mammalian cells. *Proc. Natl. Acad. Sci. USA* 77: 2163-2167.

299. Schlegel, R.A. and Rechsteiner, M.C. (1975). Microinjection of thymidine kinase and bovine serum albumin into mammalian cells by fusion with red blood cells. *Cell* 5: 371-376.

300. Schlegel, R.A. and Rechsteiner, M.C. (1978). Red cell-mediated microinjection of macromolecules into mammalian cells. *Methods Cell Biol.* 20: 341-354.

301. Schlessinger, J., Schechter, Y., Willingham, M.C. and Pastan, I. (1978). Direct visualization of binding, aggregation, and internalization of insulin and epidermal growth factor on living fibroblastic cells. *Proc. Natl. Acad. Sci. USA* 75: 2659-2663.

302. Schlichtkrull, J. (1956a). Insulin crystals. I. The minimum mole fraction of metal in insulin crystals prepared with Zn^{++}, Ca^{++}, Co^{++}, Ni^{++}, Cu^{++}, Mn^{++} or Fe^{++}. *Acta Chem. Scand.* 10: 1455-1458.

303. Schlichtkrull, J. (1956b). Insulin crystals. II. Shape of rhomboidal zinc-insulin crystals in relation to species and crystallization media. *Acta Chem. Scand.* 10: 1459-1464.

304. Seidah, N.G. and Chretien, M. (1983). Application of high-performance liquid chromatography to characterize radiolabeled peptides for radioimmunoassay, biosynthesis, and microsequence studies of polypeptide hormones. *Methods in Enzymology* 92: 292-309.

305. Selden, R.F., Skoskiewicz, M.J., Howie, K.B., Russell, P.S. and Goodman, H.M. (1987). Implantation of genetically engineered fibroblasts into mice: implications for gene therapy. *Science* 236: 714-718.

306. Serpersu, E.H., Kinosita, K.Jr. and Tsong, T.Y. (1985). Reversible and irreversible modification of erythrocyte membrane permeability by electric field. *Biochim. Biophys. Acta* 812: 779-785.

307. Sheehan, J.C. and Hess, G.P. (1955). A new method of forming peptide bonds. *J. Amer. Chem. Soc.* 77: 1067-1068.

308. Shestakov, S.V. and Khyen, N.T. (1970). Evidence for genetic transformation in the blue-green alga *Anacystis nidulans*. *Mol. Gen. Genet.* 107: 372-375.

309. Shigekawa, K. and Dower, W.J. (1988). Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells. *Biotechniques* 6: 742-751.

310. Shih, C. and Weinberg, R.A. (1982). Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* 29: 161-169.

311. Shimotohno, K. and Temin, H.M. (1981). Formation of infectious progeny virus after insertion of Herpes Simplex thymidine kinase gene into DNA of an avian retrovirus. *Cell* 26: 67-77.

312. Simon, J., Freychet, P., Rosselin, G. and de Meyts, P. (1977). Enhanced binding affinity of chicken insulin in rat liver membranes or human lymphocytes: relationship to the kinetic properties of the hormone-receptor interaction. *Endocrinology* 100: 115-121.

313. Simons, J.P., McClenaghan, M. and Clark, A.J. (1987). Alteration of the quality of milk by expression of sheep ß-lactoglobulin in transgenic mice. *Nature* 328: 530-532.

314. Singh, A., Ursic, D. and Davies, J. (1979). Phenotypic suppression and misreading in *Saccharomyces cerevisiae*. *Nature* 277: 146-148.

315. Singh, H., Sen, R., Baltimore, D. and Sharp, P.A. (1986). A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature* 319: 154-158.

316. Sirotnak, F.M. and Hutchinson, D.J. (1959). Absorption of deoxyribonucleic acid by mouse lymphoma cells. *Biochim. Biophys. Acta* 36: 246-248.

317. Smirnov, V.N., Domogatsky, S.P., Dolgov, V.V. and Hvatov, V.B. (1986). Carrier-directed targeting of liposomes and erythrocytes to denuded areas of vessel walls. *Proc. Natl. Acad. Sci. USA* 83: 6603-6607.

318. Smith, L.F. (1972). Amino acid sequences of insulins. *Diabetes* 21, Suppl. 2: 457-460.

319. Sompayrac, L.M. and Danna, K.J. (1981). Efficient infection of monkey cells with DNA of simian virus 40. *Proc. Natl. Acad. Sci. USA* 78: 7575-7578.

320. Sonne, O. (1988). Receptor-mediated endocytosis and degradation of insulin. *Physiol. Rev.* 68: 1129-1196.

321. Sonne, O., Linde, S., Larsen, T.R. and Gliemann, J. (1983). Monoiodoinsulin labelled in tyrosine residue 16 or 26 of the B-chain or 19 of the A-chain. II. Characterization of the kinetic binding constants and determination of the biological potency. *Hoppe-Seyler's Z. Physiol. Chem.* 364: 101-110.

322. Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517.

323. Southern, P.J. and Berg, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Genet.* 1: 327-341.

324. Srinivasan, P.R. and Lewis, W.H. (1980). Transfer of the dihydrofolate reductase gene into mammalian cells using metaphase chromosomes or purified DNA. In *Introduction of Macromolecules into Viable Mammalian Cells*, ed. Baserga, R., Croce, C. and Rovera, G. The Wistar Symposium Series Vol.1. pp. 27-45. Alan R. Liss, Inc., New York.

325. Stacey, D.W. (1980). Behaviour of microinjected molecules and recipient cells. In *Introduction of Macromolecules into Viable Mammalian Cells*, ed. Baserga, R., Croce, C. and Rovera, G. The Wistar Symposium Series Vol.1. pp. 125-134. Alan R. Liss, Inc., New York.

326. Stahl, D.A., Pace, B., Marsh, T. and Pace, N.R. (1984). The ribonucleoprotein substrate for a ribosomal RNA-processing nuclease. *J. Biol. Chem.* 259: 11448-11453.

327. Stasiak, A., Di Capua, E. and Koller, T. (1981). Elongation of duplex DNA by recA protein. *J. Mol. Biol.* 151: 557-564.

328. Steiner, D.F., Hallund, O., Rubenstein, A., Cho, S. and Bayliss, C. (1968). Isolation and properties of proinsulin, intermediate forms and other minor components from crystalline bovine insulin. *Diabetes* 17: 725-736.

329. Stollar, B.D. (1980). The experimental induction of antibodies to nucleic acids. *Methods in Enzymology* 70: 70-85.

330. Stopper, H., Jones, H. and Zimmermann, U. (1987). Large scale transfection of mouse L-cells by electro-permeabilization. *Biochim. Biophys. Acta* 900: 38-44.

331. Strain, A. and Wyllie, A.H. (1984). The uptake and stability of simian-virus-40 DNA after calcium phosphate transfection of CV-1 cells. *Biochem. J.* 218: 475-482.

332. Straubinger, R.M., Hong, K., Friend, D.S. and Papahadjopoulos, D. (1983). Endocytosis of liposomes and intracellular fate of encapsulated molecules: encounter with a low pH compartment after internalization in coated vesicles. *Cell* 32: 1069-1080.

333. Straubinger, R.M. and Papahadjopoulos, D. (1983). Liposomes as carriers for intracellular delivery of nucleic acids. *Methods in Enzymology* 101: 512-527.

334. Stroun, J., Stroun-Guttieres, L., Rossi, J. and Stroun, M. (1962). *Compt. Rend. Acad. Sci.* 255: 781-

335. Struck, D.K., Hoekstra, D. and Pagano, R.E. (1981). Use of resonance energy transfer to monitor membrane fusion. *Biochemistry* 20: 4093-4099.

336. Sutcliffe, J.G. (1979). Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harbour Symp. Quant. Biol.* 43: 77-90.

337. Suzuki, F., Daikuhara, Y., Ono, M. and Tekeda, Y. (1972). Studies on the mode of action of insulin: properties and biological activity of an insulin-dextran complex. *Endocrinology* 90: 1220-1230.

338. Szoka, F.Jr. and Papahadjopoulos, D. (1978). Procedure for preparation of liposomes with large internal aqueous space and high capture by reverse-phase evaporation. *Proc. Natl. Acad. Sci. USA* 75: 4194-4198.

339. Szybalska, E. and Szybalski, W. (1962). Genetics of human cell lines. IV. DNA mediated heritable transformation of a biochemical trait. *Proc. Natl. Acad. Sci. USA* 48: 2026-2034.

340. Taketo, A. (1988). DNA transfection of *E. coli* by electroporation. *Biochim. Biophys. Acta* 949: 318-324.

341. Tanaka, N. (1982). Mechanism of action of aminoglycoside antibiotics. In *Aminoglycoside antibiotics*, ed. Umezawa, H. and Hooper, I.R. pp. 221-266. Springer-Verlag, Berlin.

342. Tao, W., Wilkinson, J., Stanbridge, E.J. and Berns, M.W. (1987). Direct gene transfer into human cultured cells facilitated by laser micropuncture of the cell membrane. *Proc. Natl. Acad. Sci. USA* 84: 4180-4184.

343. Terris, S and Steiner, D.F. (1976). Retention and degradation of ^{125}I-insulin by perfused livers from diabetic rats. *J. Clin. Invest.* 57: 885-896.

344. Thomas, G. and Thorne, B.A. (1988). Gene transfer techniques to study neuropeptide processing. *Ann. Rev. Physiol.* 50: 323-332.

345. Thomas, K.R. and Capecchi, M.R. (1986). Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene. *Nature* 324: 34-38.

346. Thomas, K.R., Folger, K.R. and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. *Cell* 44: 419-428.

347. Thorell, J.I. and Johansson, B.G. (1971). Enzymatic iodination of polypeptides with ^{125}I to high specific activity. *Biochim. Biophys. Acta* 251: 363-369.

348. Timkovitch, R. (1977). Detection of the stable addition of carbodiimide to proteins. *Analyt. Biochem.* 79: 135-143.

349. Tomlinson, E. and Burger, J.J. (1985). Incorporation of water-soluble drugs in albumin microspheres. *Methods in Enzymology* 112: 27-43.

350. Tsao, C.-J., Hosoi, T., Hirai, H., Saito, T., Okabe, T., Urabe, A. and Takaku, F. (1987). Protoplast-mediated gene transfer into human leukemia (K562) cells. *Int. J. Cell Cloning* 5: 220-230.

351. Tsumita, T. and Iwanaga, M. (1963). Fate of injected deoxyribonucleic acid in mice. *Nature* 198: 1088-1089.

352. Ullrich, A., Bell, J.R., Chen, E.Y., Herrera, R., Petruzzelli, L.M., Dull, T.J., Gray, A., Cousens, L., Liao, Y.-C., Tsubokawa, M., Mason, A., Seeburg, P.H., Grunfeld, C., Rosen, O.M. and Ramachandran, J. (1985). Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes. *Nature* 313: 756-761.

353. Umezawa, H. (1974). Biochemical mechanism of resistance to aminoglycoside antibiotics. *Adv. Carbohydr. Chem. Biochem.* 30: 183-225.

354. Umezawa, H. and Kondo, S. (1982). Mechanisms of resistance to aminoglycoside antibiotics. In *Aminoglycoside Antibiotics*, ed. Umezawa, H. and Hooper, I.R. pp. 267-292. Springer-Verlag, Berlin.

355. Vainstein, A., Razin, A., Graessmann, A. and Loyter, A. (1983). Fusogenic reconstituted Sendai virus envelopes as a vehicle for introducing DNA into viable mammalian cells. *Methods in Enzymology* 101: 492-512.

356. Van den Heuvel, R.M.M., van Eys, G.J.J.M., Ramaekers, F.C.S., Quax, W.J., Vree Egberts, W.T.M., Schaart, G., Cuypers, H.T.M. and Bleomendal, H. (1987). Intermediate filament formation after transfection with modified hamster vimentin and desmin genes. *J. Cell Sci.* 88: 475-482.

357. Varga, J.M. (1985). Hormone-drug conjugates. *Methods in Enzymology* 112: 259-269.

358. Varmus, H. (1988). Retroviruses. *Science* 240: 1427-1435.

359. Varshavsky, A. (1987). Electrophoretic assay for DNA-binding proteins. *Methods in Enzymology* 151: 551-565.

360. Varshavsky, A., Bakayev, V.V. and Georgiov, G.P. (1976). Heterogeneity of chromatin subunits *in vitro* and location of histone H1. *Nuc. Acids Res.* 3: 477-492.

361. Vestweber, D. and Schatz, G. (1989). DNA-protein conjugates can enter mitochondria via the protein import pathway. *Nature* 338: 170-172.

362. Vigneri, R., Golfine, I.D., Wong, K.Y., Smith, G.J. and Pezzino, V. (1978). The nuclear envelope. The major site of insulin binding in rat liver nuclei. *J. Biol. Chem.* 253: 2098-2103.

363. Waelbroeck, M. (1982). The pH dependence of insulin binding. A quantitative study. *J. Biol. Chem.* 257: 8284-8291.

364. Waelbroeck, M., van Obberghen, E. and de Meyts, P. (1979). Thermodynamics of the interaction of insulin with its receptors. *J. Biol. Chem.* 254: 7736-7740.

365. Wagman, G.H., Marquez, J.A., Bailey, J.V., Cooper, D., Weinstein, J., Tkach, R. and Danieels, P.J.L. (1972). Chromatographic separation of some minor components of the gentamicin complex. *J. Chromatog.* 70: 171-173.

366. Wagner, M.J., Sharp, J.A. and Summers, W.C. (1981). Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. *Proc. Natl. Acad. Sci. USA* 78: 1441-1445.

367. Wahl, G.M., Meinkoth, J.L. and Kimmel, A.R. (1987). Northern and southern blots. *Methods in Enzymology* 152: 572-581.

368. Weber, I.T. and Steitz, T.A. (1984). A model for the non-specific binding of catabolite gene activator protein to DNA. *Nuc. Acids Res.* 12: 8475-8487.

369. Weber, K. and Osborn, M. (1969). The reliability of molecular weight determinations by dodecyl sulphate-polyacrylamide gel electrophoresis. *J. Biol. Chem.* 244: 4406-4412.

370. Wei, C.-M., Gibson, M., Spear, P.G. and Scolnick, E.M. (1981). Construction and isolation of a transmissible retrovirus containing the *src* gene of Harvey murine sarcoma virus and the thymidine kinase gene of Herpes Simplex Virus Type 1. *J. Virol.* 39: 935-944.

371. Weinstein, M.J., Leudemann, G.M., Oden, E.M. and Wagman, H. (1964). Gentamicin, a new broad-spectrum antibiotic complex. *Antimicrob. Agents Chemother.* 1963: 1-7.

372. Weintraub, H., Cheng, P.F. and Conrad, K. (1986). Expression of transfected DNA depends upon DNA topology. *Cell* 46: 115-122.

373. West, S.C. (1988). Protein-DNA interactions in genetic recombination. *Trends Genet.* 4: 8-13.

374. Wharton, R.P. and Ptashne, M. (1985). Changing the binding specificity of a repressor by redesigning an α-helix. *Nature* 316: 601-605.

375. Wiberg, F.C., Sunnerhagen, P., Kaltoft, K., Zeuthen, J. and Bjursell, G. (1983). Replication and expression in mammalian cells of transfected DNA; description of an improved erythrocyte ghost fusion technique. *Nuc. Acids Res.* 11: 7287-7302.

376. Wiegand, R., Weber, G., Zimmermann, K., Monajembashi, S., Wolfrum, J. and Greulich, K.-O. (1987). Laser-induced fusion of mammalian cells and plant protoplasts. *J. Cell Sci.* 88: 145-149.

377. Wigler, M.S., Siverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.C. and Axel, R. (1977). Transfer of purified herpesvirus thymidine kinase to cultured mouse cells. *Cell* 11: 223-232.

378. Wilchek, M., Oka, T. and Topper, Y.J. (1975). Structure of a soluble super-active insulin is revealed by the nature of the complex between cyanogen-bromide-activated Sepharose and amines. *Proc. Natl. Acad. Sci. USA* 72: 1055-1058.

379. Wilhelm, J.M., Pettitt, S.E. and Jessop, J.J. (1978a). Aminoglycoside antibiotics and eukaryotic protein synthesis: structure-function relationships in the stimulation of misreading with a wheat embryo system. *Biochemistry* 17: 1143-1149.

380. Wilhelm, J.M., Jessop, J.J. and Pettitt, S.E. (1978b). Aminoglycoside antibiotics and eukaryotic protein synthesis: stimulation errors in the translation of natural messengers in extracts of cultured human cells. *Biochemistry* 17: 1149-1153.

381. Williams, D.A., Lemischka, I.R., Nathan, D.G. and Mulligan, R.C. (1984). Introduction of new genetic material into pluripotent haematopoietic stem cells of the mouse. *Nature* 310: 476-480.

382. Willingham, M.C. and Pastan, I. (1980). The receptosome: an intermediate organelle of receptor-mediated endocytosis in cultured fibroblasts. *Cell* 21: 67-77.

383. Willis, R.C., Jolly, D.J., Miller, A.D., Plent, M.M., Esty, A.C., Anderson, P.J., Chang, H.-C., Jones, O.W., Seegmiller, J.E. and Friedmann, T. (1984). Partial phenotypic correction of human Lesch-Nyhan (hypoxanthine-guanine phosphoribosyltransferase-deficient) lymphoblasts with a transmissible retroviral vector. *J. Biol. Chem.* 259: 7842-7849.

384. Willmitzer, L. (1988). The use of transgenic plants to study gene expression. *Trends Genet.* 4: 13-18.

385. Wilson, J.M., Birinyi, L.K., Salomon, R.N., Libby, P., Callow, A.D. and Mulligan, R.C. (1989). Implantation of vascular grafts lined with genetically modified endothelial cells. *Science* 244: 1344-1346.

386. Winterbourne, D.J., Thomas, S., Hermon-Taylor, J., Hussain, I. and Johnstone, A.P. (1988). Electric shock-mediated transfection of cells. Characterization and optimization of electrical parameters. *Biochem. J.* 251: 427-434.

387. Wu, G.Y. and Wu, C.H. (1987). Receptor-mediated *in vitro* gene transformation by a soluble DNA carrier system. *J. Biol. Chem.* 262: 4429-4432.

388. Wu, G.Y. and Wu, C.H. (1988a). Evidence for targeted gene delivery to HepG2 hepatoma cells *in vitro*. *Biochemistry* 27: 887-892.

389. Wu, G.Y. and Wu, C.H. (1988b). Receptor-mediated gene delivery and expression *in vivo*. *J. Biol. Chem.* 263: 14621-14624.

390. Xiong, C., Levis, R., Shen, P., Schlesinger, S., Rice, C.M. and Huang, H.V. (1989). Sindbis virus: an efficient broad host range vector for gene expression in animal cells. *Science* 243: 1188-1191.

391. Yapel, A.F. (1985). Albumin microspheres: heat and chemical stabilization. *Methods in Enzymology* 112: 3-18.

392. Zeleznik, A.J. and Roth, J. (1977). Demonstration of the insulin receptor *in vivo* in rabbits and its possible role as a reservoir for the plasma hormone. *J. Clin. Invest.* 61: 1363-1374.

393. Zimmermann, U. (1982). Electric field-mediated fusion and related electrical phenomena. *Biochim. Biophys. Acta* 694: 227-277.

394. Zwiebel, J.A., Freeman, S.M., Kantoff, P.W., Cornetta, K., Ryan, U.S. and French Anderson, W. (1989). High-level recombinant gene expression in rabbit endothelial cells transduced by retroviral vectors. *Science* 243: 220-222.

APPENDIX I

NUCLEOTIDE SEQUENCE OF THE EXPRESSION VECTOR ptkNEO

The vector ptkNEO was previously called pNEO3.

The sequence was compiled from known details of Barbara Wold's original construction provided by Dr Shirley Taylor and Nevis Fregien, in conjunction with the respective published sequences of pBR322 (Sutcliffe, 1979), the Herpes Simplex Virus Type 1 *tk* gene (Wagner *et al*, 1981) and the Tn5 *neo* gene (in Gorman, 1985).

sequence		base position	size
PvuII*-BamHI	pBR322 *amp*^{R} and ori	1-2653	(2653bp)
BamHI-BglII	HSV1 *tk* 5′ proximal to *neo*	2654-3117	(464bp)
BglII-SalI	Tn5 *neo*	3118-4287	(1170bp)
SmaI*-PvuII	HSV1 *tk* 3′ proximal to *neo*	4288-4913	(626bp)

* Some bases have been removed in the vicinity of ligation sites: PvuII and SmaI sites destroyed.

***tk*/pBR322 ligation site**
PvuII (17bp of the pBR322 sequence removed)

```
        10          20          30          40          50
GTGATGACGG  TGAAAACCTC  TGACACATGC  AGCTCCCGGA  GACGGTCACA

        60          70          80          90         100
GCTTGTCTGT  AAGCGGATGC  CGGGAGCAGA  CAAGCCCGTC  AGGGCGCGTC

       110         120         130         140         150
AGCGGGTGTT  GGCGGGTGTC  GGGGCGCAGC  CATGACCCAG  TCACGTAGCG

       160         170         180         190         200
ATAGCGGAGT  GTATACTGGC  TTAACTATGC  GGCATCAGAG  CAGATTGTAC

       210         220         230         240         250
TGAGAGTGCA  CCATATGCGG  TGTGAAATAC  CGCACAGATG  CGTAAGGAGA

       260         270         280         290         300
AAATACCGCA  TCAGGCGCTC  TTCCGCTTCC  TCGCTCACTG  ACTCGCTGCG
```

```
       310        320        330        340        350
CTCGGTCGTT CGGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA

       360        370        380        390        400
TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA

       410        420        430        440        450
AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT

       460        470        480        490        500
TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA

       510        520        530        540        550
GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA GGCGTTTCCC

       560        570        580        590        600
CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT CCGACCCTGC CGCTTACCGG

       610        620        630        640        650
ATACCTGTCC GCCTTTCTCC CTTCGGGAAG CGTGGCGCTT TCTCAATGCT
                                   <ori>

       660        670        680        690        700
CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC

       710        720        730        740        750
TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA

       760        770        780        790        800
CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG

       810        820        830        840        850
CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA

       860        870        880        890        900
GAGTTCTTGA AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT

       910        920        930        940        950
TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA

       960        970        980        990       1000
GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTGG TTTTTTTGTT
```

1010 1020 1030 1040 1050
TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCTTT

1060 1070 1080 1090 1100
GATCTTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAG

1110 1120 1130 1140 1150
GGATTTTGGT CATGAGATTA TCAAAAAGGA TCTTCACCTA GATCCTTTTA
end *amp*^{R} ◀—

1160 1170 1180 1190 1200
AATTAAAAAT GAAGTTTTAA ATCAATCTAA AGTATATATG AGTAAACTTG

1210 1220 1230 1240 1250
GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC TCAGCGATCT

1260 1270 1280 1290 1300
GTCTATTTCG TTCATCCATA GTTGCCTGAC TCCCCGTCGT GTAGATAACT

1310 1320 1330 1340 1350
ACGATACGGG AGGGCTTACC ATCTGGCCCC AGTGCTGCAA TGATACCGCG

1360 1370 1380 1390 1400
AGACCCACGC TCACCGGCTC CAGATTTATC AGCAATAAAC CAGCCAGCCG

1410 1420 1430 1440 1450
GAAGGGCCGA GCGCAGAAGT GGTCCTGCAA CTTTATCCGC CTCCATCCAG

1460 1470 1480 1490 1500
TCTATTAATT GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC CAGTTAATAG

1510 1520 1530 1540 1550
TTTGCGCAAC GTTGTTGCCA TTGCTGCAGG CATCGTGGTG TCACGCTCGT
PstI ▲

1560 1570 1580 1590 1600
CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT

1610 1620 1630 1640 1650
ACATGATCCC CCATGTTGTG CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC

1660 1670 1680 1690 1700
GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG

1710 1720 1730 1740 1750
CAGCACTGCA TAATTCTCTT ACTGTCATGC CATCCGTAAG ATGCTTTTCT

1760 1770 1780 1790 1800
GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT GTATGCGGCG

1810 1820 1830 1840 1850
ACCGAGTTGC TCTTGCCCGG CGTCAACACG GGATAATACC GCGCCACATA

1860 1870 1880 1890 1900
GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGGCGAAAA

1910 1920 1930 1940 1950
CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT AACCCACTCG

1960 1970 1980 1990 2000
TGCACCCAAC TGATCTTCAG CATCTTTTAC TTTCACCAGC GTTTCTTGGGT

2010 2020 2030 2040 2050
GAGCAAAAAC AGGAAGGCAA AATGCCGCAA AAAAGGGAAT AAGGGCGACA
← **start** *amp*^{R}

2060 2070 2080 2090 2100
CGGAAATGTT GAATACTCAT ACTCTTCCTT TTTCAATATT ATTGAAGCAT

2110 2120 2130 2140 2150
TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTTGAA TGTATTTAGA

2160 2170 2180 2190 2200
AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCCGAAA AGTGCCACCT

2210 2220 2230 2240 2250
GACGTCTAAG AAACCATTAT TATCATGACA TTAACCTATA AAAATAGGCG

2260 2270 2280 2290 2300
TATCACGAGG CCCTTTCGTC TTCAAGAATT CTCATGTTTG ACAGCTTATC

2310 2320 2330 2340 2350
ATCGATAAGC TTTAATGCGG TAGTTTATCA CAGTTAAATT GCTAACGCAG

2360 2370 2380 2390 2400
TCAGGCACCG TGTATGAAAT CTAACAATGC GCTCATCGTC ATCCTCGGCA

```
      2410        2420        2430        2440        2450
CCGTCACCCT  GGATGCTGTA  GGCATAGGCT  TGGTTATGCC  GGTACTGCCG

      2460        2470        2480        2490        2500
GGCCTCTTGC  GGGATATCGT  CCATTCCGAC  AGCATCGCCA  GTCACTATGG

      2510        2520        2530        2540        2550
CGTGCTGCTA  GCGCTATATG  CGTTGATGCA  ATTTCTATGC  GCACCCGTTC

      2560        2570        2580        2590        2600
TCGGAGCACT  GTCCGACCGC  TTTGGCCGCC  GCCCAGTCCT  GCTCGCTTCG

      2610        2620        2630        2640        2650
CTACTTGGAG  CCACTATCGA  CTACGCGATC  ATGGCGACCA  CACCCGTCCT

      2660        2670        2680        2690        2700
GTGGATCCGG  GTCCTAGGCT  CCATGGGGAC  CGTATACGTG  GACAGGCTCT
          pBR322/tk ligation site
BamHI
      2710        2720        2730        2740        2750
GGAGCATCGC  ACGACTGCGT  GATATTACCG  GAGACCTTCT  GCGGGACGAG

      2760        2770        2780        2790        2800
CCGGGTCACG  CGGCTGACGG  AGCGTCCGTT  GGGCGACAAA  CACCAGGACG

      2810        2820        2830        2840        2850
GGGCACAGGT  ACACTATCTT  GTCACCCGGA  GCGCGAGGGA  CTGCAGGAGC
                                                    PstI
      2860        2870        2880        2890        2900
TTCAGGGAGT  GGCGCAGCTG  CTTCATCCCC  GTGGCCCGTT  GCTCGCGTTT

      2910        2920        2930        2940        2950
GCTGGCGGTG  TCCCCGGAAG  AAATATATTT  GCATGTCTTT  AGTTCTATGA

      2960        2970        2980        2990        3000
TGACACAAAC  CCCGCCCAGC  GTCTTGTCAT  TGGCGAATTC  GAACACGCAG
                                  possible control sequence
                                    (Benoist et al, 1980)
      3010        3020        3030        3040        3050
ATGCAGTCGG  GGCGGCGCGG  TCCCAGGTCC  ACTTCGCATA TTAAGGTGAC
                                  Goldberg-Hogness 'TATA' box

      3060        3070        3080        3090        3100
GCGTGTGGCC  TCGAACACCG  AGCGACCCTG  CAGCGACCCG  CTTAACAGCG
                                        PstI
```

3110 3120 3130 3140 3150
TCAACAGCGT GCCGCAAGAT CTGATCAAGA GACAGGATGA GGATCGTTTC

tk/neo ligation site

BglII

3160 3170 3180 3190 3200
GCATGATTGA ACAAGATGGA TTGCACGCAG GTTCTCCGGC CGCTTGGGTG

3210 3220 3230 3240 3250
GAGAGGCTAT TCGGCTATGA CTGGGCACAA CAGACAATCG GCTGCTCTGA

3260 3270 3280 3290 3300
TGCCGCCGTG TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT CTTTTTGTCA

3310 3320 3330 3340 3350
AGACCGACCT GTCCGGTGCC CTGAATGAAC TGCAGGACGA GGCAGCGCGG

PstI

3360 3370 3380 3390 3400
CTATCGTGGC TGGCCACGAC GGGCGTTCCT TGCGCAGCTG TGCTCGACGT

3410 3420 3430 3440 3450
TGTCACTGAA GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA GTGCCGGGGC

3460 3470 3480 3490 3500
AGGATCTCCT GTCATCTCAC CTTGCTCCTG CCGAGAAAGT ATCCATCATG

3510 3520 3530 3540 3550
GCTGATGCAA TGCGGCGGCT GCATACGCTT GATCCGGCTA CCTGCCCATT

3560 3570 3580 3590 3600
CGACCACCAA GCGAAACATC GCATCGAGCG AGCACGTACT CGGATGGAAG

3610 3620 3630 3640 3650
CCGGTCTTGT CGATCAGGAT GATCTGGACG AAGAGCATCA GGGGCTCGCG

3660 3670 3680 3690 3700
CCAGCCGAAC TGTTCGCCAG GCTCAAGGCG CGCATGCCCG ACGGCGAGGA

3710 3720 3730 3740 3750
TCTCGTCGTG ACCCATGGCG ATGCCTGCTT GCCGAATATC ATGGTGGAAA

3760 3770 3780 3790 3800
ATGGCCGCTT TTCTGGATTC ATCGACTGTG GCCGGCTGGG TGTGGCGGAC

3810 3820 3830 3840 3850
CGCTATCAGG ACATAGCGTT GGCTACCCGT GATATTGCTG AAGAGCTTGG

3860 3870 3880 3890 3900
CGGCGAATGG GCTGACCGCT TCCTCGTGCT TTACGGTATC GCCGCTCCCG

3910 3920 3930 3940 3950
ATTCGCAGCG CATCGCCTTC TATCGCCTTC TTGACGAGTT CTTCTGAGCG

3960 3970 3980 3990 4000
GGACTCTGGG GTTCGAAATG ACCGACCAAG CGACGCCCAA CCTGCCATCA

4010 4020 4030 4040 4050
CGAGATTTCG ATTCCACCGC CGCCTTCTAT GAAAGGTTGG GCTTCGGAAT

4060 4070 4080 4090 4100
CGTTTTCCGG GACGCCGGCT GGATGATCCT CCAGCGCGGG GATCTCATGC

4110 4120 4130 4140 4150
TGGAGTTCTT CGCCCACCCC GGGCTCGATC CCCTCGCGAG TTGGTTCAGC

4160 4170 4180 4190 4200
TGCTGCCTGA GGCTGGACGA CCTCGCGGAG TTCTACCGGC AGTGCAAATC

4210 4220 4230 4240 4250
CGTCGGCATC CAGGAAACCA GCAGCGGCTA TCCGCGCATC CATGCCCCCG

4260 4270 4280 4290 4300
AACTGCAGGA GTGGGGAGGC ACGATGGCCG CTTTGGCGGG AGATGGGGGA

▲ *neo/tk* **ligation site** ▲
PstI (some bases removed) **SmaI**

4310 4320 4330 4340 4350
GGCTAACTGA AACACGGAAG GAGACAATAC CGGAAGGAAC CCGCGCTATG

4360 4370 4380 4390 4400
ACGGC**AATAA AA**AGACAG**AA TAAA**ACGCAC GGGTGTTGGG TCGTTTGTTC
polyadenylation signals

4410 4420 4430 4440 4450
ATAAACGCGG GGTTCGGTCC CAGGGCTGGC ACTCTGTCGA TACCCCACCG

4460 4470 4480 4490 4500
AGACCCCATT GGGGCCAATA CGCCCGCGTT TCTTCCTTTT CCCCACCCCA

```
      4510        4520        4530        4540        4550
ACCCCCAAGT  TCGGGTGAAG  GCCCAGGGCT  CGCAGCCAAC  GTCGGGGCGG

      4560        4570        4580        4590        4600
CAAGCCCGCC  ATAGCCACGG  GCCCCGTGGG  TTAGGGACGG  GGTCCCCCAT

      4610        4620        4630        4640        4650
GGGGAATGGT  TTATGGTTCG  TGGGGGTTAT  TCTTTTGGGC  GTTGCGTGGG

      4660        4670        4680        4690        4700
GTCAGGTCCA  CGACTGGACT  GAGCAGACAG  ACCCATGGTT  TTTGGATGGC

      4710        4720        4730        4740        4750
CTGGGCATGG  ACCGCATGTA  CTGGCGCGAC  ACGAACACCG  GGCGTCTGTG

      4760        4770        4780        4790        4800
GCTGCCAAAC  ACCCCCGACC  CCCAAAAACC  ACCGCGCGGA  TTTCTGGCGC

      4810        4820        4830        4840        4850
CGCCGGACGA  ACTAAACCTG  ACTACGGCAT  CTCTGCCCCT  TCTTCGCTGG

      4860        4870        4880        4890        4900
TACGAGGAGC  GCTTTTGTTT  TGTATTGGTC  ACCACGGCCG  AGTTTCCGCG

      4910        4920
GGACCCCGGC  CAG
```

PvuII

tk/pBR322 ligation site

APPENDIX II

PUBLICATIONS

1. Huckett, B., Gordhan, H., Hawtrey, R., Moodley, N., Ariatti, M. and Hawtrey, A. (1986). Binding of DNA to albumin and transferrin modified by treatment with water-soluble carbodiimides. *Biochem. Pharmacol.* 35: 1249-1257.

   (Reprint attached.)

2. Huckett, B., Ariatti, M. and Hawtrey, A. Evidence for targeted gene transfer by receptor-mediated endocytosis: stable expression following insulin-directed entry of *neo* into HepG2 cells.

   (Manuscript submitted to *Biochemical Pharmacology*, November, 1989. Copy attached.)

Biochemical Pharmacology, Vol. 35, No. 8, pp. 1249–1257, 1986.
Printed in Great Britain.

0006-2952/86 $3.00 + 0.00

BINDING OF DNA TO ALBUMIN AND TRANSFERRIN MODIFIED BY TREATMENT WITH WATER-SOLUBLE CARBODIIMIDES

BARBARA HUCKETT, HASHA GORDHAN, RICHARD HAWTREY, NIRINJINI MOODLEY, MARIO ARIATTI and ARTHUR HAWTREY*

Department of Biochemistry, University of Durban-Westville, Durban 4000, South Africa

(*Received* 29 *July* 1985; *accepted* 18 *October* 1985)

Abstract—*N*-Acylurea derivatives of albumin and transferrin prepared with the water-soluble carbodiimides *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide and *N*-ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide have been found to bind different types of DNA. The two proteins were reacted with varying amounts of carbodiimide in water at pH 5.5 for 36–60 hr at 20°, and then purified. In the case of iron-loaded transferrin, reactions with carbodiimides were in phosphate-buffered saline (pH 7.5) to prevent loss of iron from the protein. [^{3}H]*N*-Ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide was used for the determination of covalently attached *N*-acylurea groups in the modified proteins, and gel electrophoresis for changes in charge and possible aggregation through cross-linking. Binding of DNA to *N*-acylurea proteins was studied by means of agarose gel electrophoresis and nitrocellulose filter binding. *N*-Acylurea albumin and *N*-acylurea transferrin at low concentrations retarded the migration of λ-PstI restriction fragments, pBR322 plasmid and M13 mp8 single-standed DNA on agarose gels, while at higher concentrations of modified protein the *N*-acylurea protein–DNA complexes were unable to enter the gel. Nitrocellulose filter assays showed that binding of pBR322 DNA and calf thymus DNA to *N*-acylurea proteins is rapid and dependent on protein concentration and the ionic strength of the medium. *N*-Acylurea albumins prepared with each of the two carbodiimides gave comparable plots for DNA bound versus protein concentration. On the other hand, binding of DNA by *N*-acylurea transferrins differed according to the carbodiimide used in the synthesis. *N*-Acylurea CDI-transferrin (prepared with tertiary carbodiimide) was less effective than either of the two *N*-acylurea albumins in binding DNA. In contrast with these results, *N*-acylurea Me$^+$CDI-transferrin (prepared with quaternary carbodiimide) was far more effective in binding DNA and in this respect was similar to the *N*-acylurea albumins. On the basis of experiments in which *N*-acylurea protein–DNA complexes were treated with heparin, two types of binding could be distinguished. These were (i) a weak binding occurring in the initial stages of interaction and (ii) a tight binding which developed on further incubation of the complexes. These studies show that binding of DNA by *N*-acylurea proteins is a reversible process dependent on ionic strength; interaction appears to be electrostatic in nature, although other forms of binding might be involved. The possible use of *N*-acylurea proteins for DNA transfer into cell systems is discussed.

A well established procedure for raising antibodies to drugs such as 9-β-D-arabinofuranosylcytosine (ara-C) for the purpose of radioimmunoassay is to use as the immunogen ara-C covalently conjugated to a protein such as albumin. In order to obtain specificity for the arabinose moiety of the ara-C, the 5'-*O*-succinyl derivative of the nucleoside is coupled to lysine side chains of the albumin by means of the water-soluble carbodiimide *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide (CDI)† [1, 2]. Carbodiimide-mediated coupling is a general procedure for linking many different types of hapten to carrier proteins [3] and is popular because of its simplicity of operation and efficiency under mild conditions.

During experiments concerned with the preparation of nucleoside–albumin conjugates using CDI as the coupling agent, we also reacted the protein with the carbodiimide in the absence of nucleoside to obtain the *N*-acylurea substituted protein [4–6]. This was used for certain studies during the course of which we observed that the *N*-acylurea protein was able to bind different types of DNA. This observation is interesting and has led us to investigate the mechanism of the interaction between carbodiimide-modified proteins and DNA and also to enquire whether we could modify proteins such as transferrin, asialo α_1-acid glycoprotein and others to bind DNA without altering their interaction with specific cell surface receptors. Receptor-mediated endocytosis of the protein (ligand)–DNA complexes might be possible, thus allowing selective transfer of DNA into cells [7, 8].

This work describes (i) initial experiments on the modification of bovine serum albumin and human serum transferrin by the water-soluble carbodiimides *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide HCl (CDI) and *N*-ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide (Me$^+$CDI), a CDI derivative which contains a positively charged quaternary nitrogen (Fig. 2), and (ii) a study of the interaction of these modified proteins with different classes of DNA.

* To whom correspondence should be addressed.

† *Abbreviations:* CDI, *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide; Me$^+$CDI, *N*-ethyl-*N'*-(3-trimethylpropylammonium) carbodiimide iodide; PBS, phosphate-buffered saline; SSC, saline sodium citrate.

MATERIALS AND METHODS

Materials. Bovine serum albumin, human serum transferrin, heparin, calf thymus DNA and DNase 1 (EC 3.1.21.1) were obtained from the Sigma Chemical Co., St Louis, MO. Restriction enzyme PstI (EC 3.1.23.31), pBR322 DNA, λ DNA and DNA polymerase 1 (EC 2.7.7.7) were from Boehringer–Mannheim. [^{3}H]Methyl iodide (specific radioactivity, 85 Ci/mmol), [^{3}H]dTTP (specific radioactivity, 97 Ci/mmol) and M13 mp8 single-stranded DNA were from Amersham, U.K. *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide HCl (CDI) was from Merck. Agarose (ultra pure grade), acrylamide and bis-acrylamide were obtained from BioRad. DEAE Sephacel and Sephadex G-50 were from Pharmacia. All other reagents were of Analar grade.

Sheared calf thymus DNA. Calf thymus DNA (approximately 12 kb in size) was prepared by passing a solution of DNA (0.1 mg/ml) in 0.1 SSC (SSC, 0.15 M NaCl, 0.015 M trisodium citrate) through a 25G-Yale syringe needle six times.

[^{3}H]*calf thymus and pBR322 DNA*. These were labelled by the nick translation procedure [9] and the DNA purified by the spun-column method [10] using Sephadex G-50.

[^{3}H]calf thymus DNA, 6.5×10^6 c.p.m./μg. [^{3}H]-pBR322 DNA, 7.6×10^6 c.p.m./μg.

Protein determination. Protein was determined by the method of Lowry [11] using either bovine serum albumin or human serum transferrin as standard.

[^{3}H]N-*Ethyl*-N'-(3-*trimethylpropylammonium*)-*carbodiimide iodide*. The quaternary carbodiimide was prepared essentially according to the method of Kopczynski and Babior [12], with minor changes, from the hydrochloride salt of CDI (500 mg, 2.62 mmol) and methyl iodide (0.71 g, 5 mmol) containing 250 μCi [^{3}H]methyl iodide. The hygroscopic product (520 mg, 1.76 mmol) with a specific radioactivity of 44 μCi/mmol was stored in the dark over P_2O_5 at −15°.

Modification of proteins with carbodiimides. Albumin and transferrin respectively were treated with either CDI or Me$^+$CDI at various carbodiimide to protein mole ratios which ranged from 10:1 to 2000:1. A similar reaction procedure was followed in each case. The following procedure is for a typical reaction of CDI with transferrin at a mole ratio of 500:1.

N-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide HCl (15 mg, 80 μmol) in 1.5 ml of water was added to an aqueous solution (1.3 ml) of transferrin (12 mg, 0.16 μmol) and the clear reaction mixture carefully adjusted to pH 5.5 by the addition of dilute HCl. The mixture was allowed to stand at room temperature (18–20°) for 48–60 hr in the dark and then dialysed exhaustively against 0.05 M NaCl containing 0.3 mM EDTA (pH 7.0) at 5°. Aliquots of the dialysed product, *N*-acylurea CDI-transferrin, were stored at −15°.

Purification of the *N*-acylurea protein conjugates can be achieved by chromatography on small columns of DEAE-cellulose (Sephacel) of dimensions 1.4 × 12 cm using 0.05 M NaCl containing 0.01 M Tris–HCl (pH 7.0) as the eluent. *N*-Acylurea proteins are not retained on the column under these conditions in contrast with the unmodified proteins.

N-*Acylurea* Me$^+$*CDI*-(Fe^{3+})*transferrin*. Transferrin (4 mg, 0.04 μmol) saturated with Fe^{3+} ion [13] in 0.6 ml of phosphate-buffered saline (PBS) (pH 7.5) was added to *N*-ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide (7.35 mg, 24.7 μmol) dissolved in 0.8 ml water. The reaction mixture (pH 7.5) was incubated at 20° for 36 hr in the dark. It was then dialysed exhaustively against 0.1 M NaCl at 5° and stored at −15°.

In an alternative procedure, iron-free transferrin was reacted with *N*-ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide in PBS (pH 7.5) under identical conditions to those described above for the iron-saturated protein. It was dialysed against 0.1 M NaCl and then treated with ferric citrate before storage. In each of these alternative procedures the reaction mixtures represent a carbodiimide to protein mole ratio of 500:1.

Analysis for iron was by the α-dipyridyl method [14] and spectra (λ_{max} at 465 nm).

Number of N-*acylurea groups attached to proteins*. For this determination, transferrin or albumin was reacted with [^{3}H]*N*-ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide exactly as described above for the various *N*-acylurea protein preparations. Reaction was either in water at pH 5.5 or PBS at pH 7.5. For radioactive counting, aliquots of the modified transferrins were precipitated with an equal volume of cold 10% (w/v) TCA, filtered on Whatman GF/C filters, washed with 25 ml of cold 5% (w/v) TCA, dried and counted in scintillation fluid (Beckman HP/b). In the case of the modified albumin preparations, precipitation was with cold 60% (w/v) TCA and subsequent washing with cold 30% (w/v) TCA; all other aspects of treatment were the same as for the transferrins.

Agarose gel electrophoresis of N-*acylurea protein–DNA complexes*. *N*-Acylurea proteins were incubated for 20 min at 18–20° with one of the following: λ-PstI restriction fragments, pBR322 DNA or M13 mp8 DNA. Each reaction mixture contained in a final volume of 10 μl: λ-PstI (0.48 μg DNA) or pBR322 (0.33 μg DNA) or M13 mp8 (0.28 μg DNA), a variable concentration of *N*-acylurea protein, 6.6 mM Tris–HCl (pH 7.0), 3 mM sodium phosphate, 100 mM NaCl, 2 mM $MgCl_2$ and 0.1 mM EDTA. The incubated samples (*N*-acylurea protein–DNA complexes), as well as DNA controls, were mixed with a stop solution containing sucrose, urea, bromophenol blue and EDTA, then run in 1.2% agarose gels with a buffer containing 36 mM Tris–HCl, 30 mM sodium phosphate and 10 mM EDTA (final pH 7.5). Gels were stained with ethidium bromide (1.5 μg/ml).

Agarose gel electrophoresis of proteins. Proteins and *N*-acylurea proteins were mixed with stop solution and run in 1.2% agarose gels as described above. Gels were stained with Coomassie Brilliant Blue in methanol–acetic acid–water (5:1:5, v/v/v).

SDS–polyacrylamide gel electrophoresis. Proteins and *N*-acylurea proteins were analysed by SDS–polyacrylamide gel electrophoresis according to Laemmli [15].

Nitrocellulose filter binding assay. Binding assays

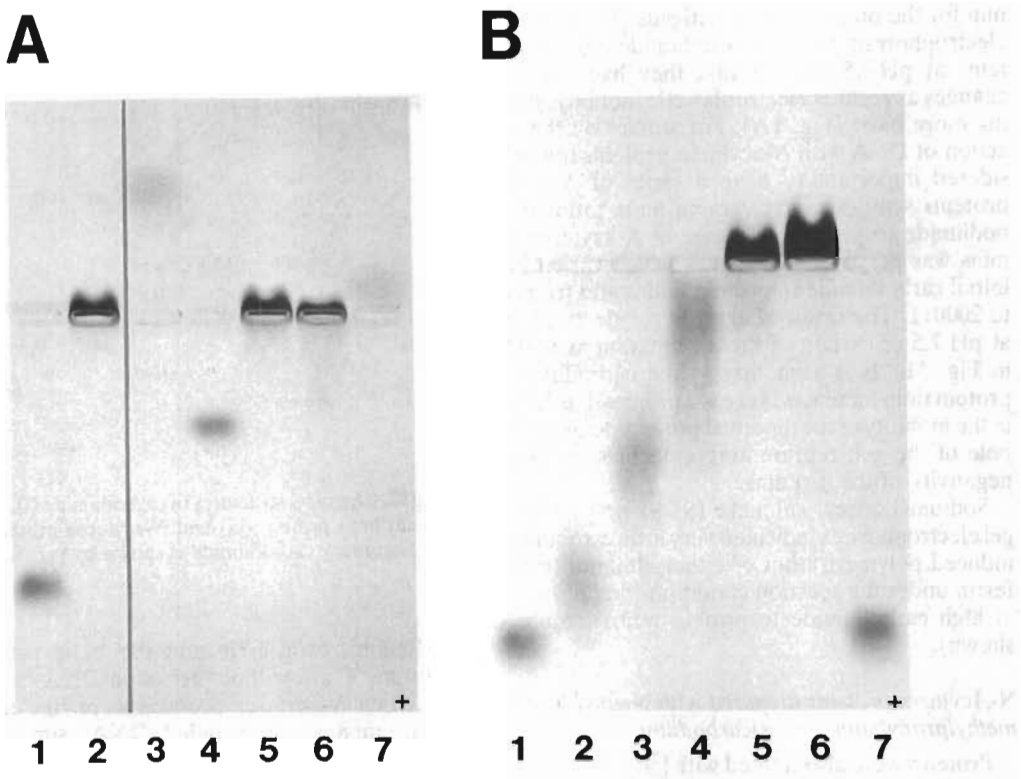


Fig. 1. Agarose gel electrophoresis of proteins and proteins modified with *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide. Protein load, 10–15 μg per well. (A) *N*-Acylurea proteins prepared at a carbodiimide to protein mole ratio of 500:1. 1, albumin; 2, *N*-acylurea CDI-albumin; 3, cytochrome C; 4, transferrin; 5 and 6, different amounts of *N*-acylurea CDI-transferrin and 7, RNase A. (B) *N*-Acylurea albumins prepared at a range of carbodiimide to protein mole ratios. 1, albumin; 2–6, *N*-acylurea albumins prepared at ratios of 10:1, 50:1, 100:1, 250:1 and 500:1 respectively; 7, albumin.

were carried out in the following manner: Each reaction mixture contained the following in a final volume of 200 μl: [^{3}H]calf thymus DNA (0.01 μg, 6.5×10^4 c.p.m.) or [^{3}H]pBR322 DNA (0.01 μg, 7.6×10^4 c.p.m.), 10 mM Tris–HCl (pH 7.5), 0.05 M NaCl, 5 mM EDTA and protein as indicated in the figures. Reaction mixtures were incubated at 18–20° for 30 min, then carefully filtered through presoaked nitrocellulose filters (Millipore Type HA, 0.45 μm) and washed with 2.0 ml of buffer containing 10 mM Tris–HCl (pH 7.5), 0.05 M NaCl and 5 mM EDTA. Filters were dried, then counted in scintillation fluid.

In certain binding experiments incubations were carried out in a buffer of higher ionic strength: 10–50 mM Tris–HCl, 0.05–1.0 M NaCl and 5 mM EDTA. Washing of the nitrocellulose filters was then executed with the buffer used for the incubation.

RESULTS

N-Acylurea proteins prepared with N-*ethyl*-N'-(3-*dimethylaminopropyl)carbodiimide*

Albumin and transferrin were treated with the water-soluble carbodiimide *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide HCl at a carbodiimide to protein mole ratio of 500:1 under dilute reaction conditions at pH 5.5. These conditions are similar to those used for the coupling of small ligands to albu-

Table 1. Reaction of [^{3}H]*N*-ethyl-*N'*-(3-trimethylpropylammonium)carbodiimide iodide with albumin and transferrin

Protein	$\frac{\text{Carbodiimide (mole)}}{\text{Protein (mole)}}$ ratio	pH of reaction	Moles *N*-acylurea attached per mole protein	Moles (side chain carboxyl groups) per mole protein
Albumin	500:1	5.5	27	97
Albumin	150:1	5.5	13	97
Transferrin	500:1	5.5	20	85
Transferrin	150:1	5.5	4	85
(Fe^{3+})Transferrin*	500:1	7.5	17	85
(Fe^{3+})Transferrin†	500:1	7.5	23	85

* Transferrin was saturated with iron before carbodiimide treatment.
† Transferrin was first modified by the carbodiimide at pH 7.5 and then iron-loaded.

min for the preparation of antigens [1]. Agarose gel electrophoresis of the carbodiimide-modified proteins at pH 7.5 showed that they had undergone changes as regards electrophoretic mobility, becoming more basic (Fig. 1A). For studies on the interaction of DNA with *N*-acylurea proteins it was considered important to have a series of *N*-acylurea proteins synthesized at varying mole ratios of carbodiimide to protein. A range of *N*-acylurea albumins was prepared, therefore, which varied in the initial carbodiimide to protein mole ratio from 10:1 to 2000:1. The result of agarose gel electrophoresis at pH 7.5 of certain of these preparations is shown in Fig. 1B. It is seen that as the carbodiimide to protein ratio increases there is a progressive decrease in the mobility of the modified proteins to the positive pole of the gel, representing a decrease in electronegativity of the proteins.

Sodium dodecyl sulphate (SDS)–polyacrylamide gel electrophoresis indicated very little carbodiimide-induced polymerization of either albumin or transferrin under the reaction conditions described, even at high carbodiimide to protein ratios (results not shown).

N-*Acylurea proteins prepared with* N-*ethyl*-N′-(3-*trimethylpropylammonium*)*carbodiimide iodide*

Proteins were also treated with [^{3}H]*N*-ethyl-*N*′-(3-trimethylpropylammonium)carbodiimide iodide for the purpose of (i) introducing into the protein concerned the *N*-acylurea grouping containing a positively charged quaternary nitrogen function (Fig. 2) and (ii) allowing determination of the number of *N*-acylurea groups covalently attached to the protein after carbodiimide modification [16, 17]. Table 1 shows the number of *N*-acylurea groups covalently attached to albumin and transferrin respectively following reaction with the [^{3}H]-labelled quaternary carbodiimide at two different initial carbodiimide to protein mole ratios, 150:1 and 500:1. The values for both albumin and transferrin show that the higher the carbodiimide to protein ratio during reaction, the greater the number of *N*-acylurea groups attached to the protein product. In the case of iron-loaded transferrin treated with the carbodiimide at a 500:1 ratio (pH 7.5), the number of *N*-acylurea groups which become covalently attached is slightly different according to whether the transferrin was iron-loaded before or after carbodiimide modification. Analysis had shown that the transferrin pre-loaded with iron suffered no loss of Fe^{3+} ion on carbodiimide modification, and that the iron-free transferrin treated with Fe^{3+} ion after carbodiimide modification was able to bind the ion successfully.

Agarose gel electrophoresis of N-*acylurea protein–DNA complexes*

Results presented in Fig. 3A show the interaction and retardation of migration of λ-PstI DNA fragments by *N*-acylurea CDI-transferrin. Results presented in Fig. 3B show that *N*-acylurea CDI-derivatives of transferrin and albumin also interact with and retard the migration of pBR322 DNA. At higher concentrations of modified protein the *N*-acylurea protein–DNA complexes are unable to enter the gel.

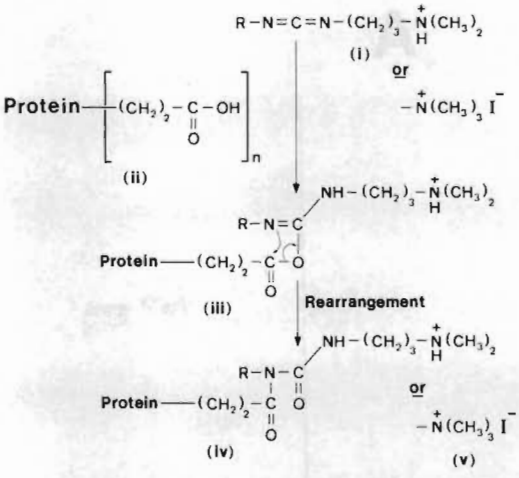


Fig. 2. Chemical structures of carbodiimide (i), protein (ii), *O*-acylurea protein (iii) and *N*-acylurea protein (iv). The quaternary carbodiimide is shown by (v). R, $—C_2H_5$.

Albumin, transferrin and the basic protein cytochrome C are without effect on DNA migration.

When *N*-acylurea proteins were run in gels with M13 mp8 single-stranded DNA, similar positive interactions were observed (results not shown).

Binding of DNA to N-*acylurea proteins using the nitrocellulose filtration method of assay*

The binding of DNA to *N*-acylurea proteins was also investigated by filtration of protein–DNA complexes through nitrocellulose filters [18–22]. The ionic strength of the incubation and washing solution was carefully controlled as the salt concentration was found to be important in the binding reactions. In these experiments we used a buffer containing 0.01 M Tris–HCl (pH 7.5), 0.05 M NaCl and 5 mM EDTA.

Figure 4 shows the binding of pBR322 DNA and calf thymus DNA to *N*-acylurea CDI-albumin at varying concentrations of protein. The two different types of DNA show similar binding characteristics up to a value of 0.05 μg protein. Thereafter, the level of calf thymus DNA binding remains constant with increasing protein concentration. Maximum binding of pBR322 DNA occurs at a value of 0.1 μg protein. Figure 4 also shows the binding of both types of DNA to *N*-acylurea Me^+CDI-albumin. Both pBR322 DNA and calf thymus DNA exhibit maximum binding with this *N*-acylurea protein at 0.02–0.05 μg protein.

Experiments on the binding of DNA by different *N*-acylurea transferrins was also carried out, and the results are shown in Fig. 5. In contrast with the binding of DNA by *N*-acylurea CDI-albumin (Fig. 4), *N*-acylurea CDI-transferrin was less efficient, requiring approximately ten times more protein (1 μg) for maximal binding. On the other hand, *N*-acylurea Me^+CDI-transferrin was able to bind both pBR322 DNA and calf thymus DNA maximally in the range 0.02–0.05 μg protein, and is therefore considerably more effective than *N*-acylurea CDI-transferrin in binding DNA. It is of interest to note that *N*-acylurea Me^+CDI-(Fe^{3+})transferrin showed

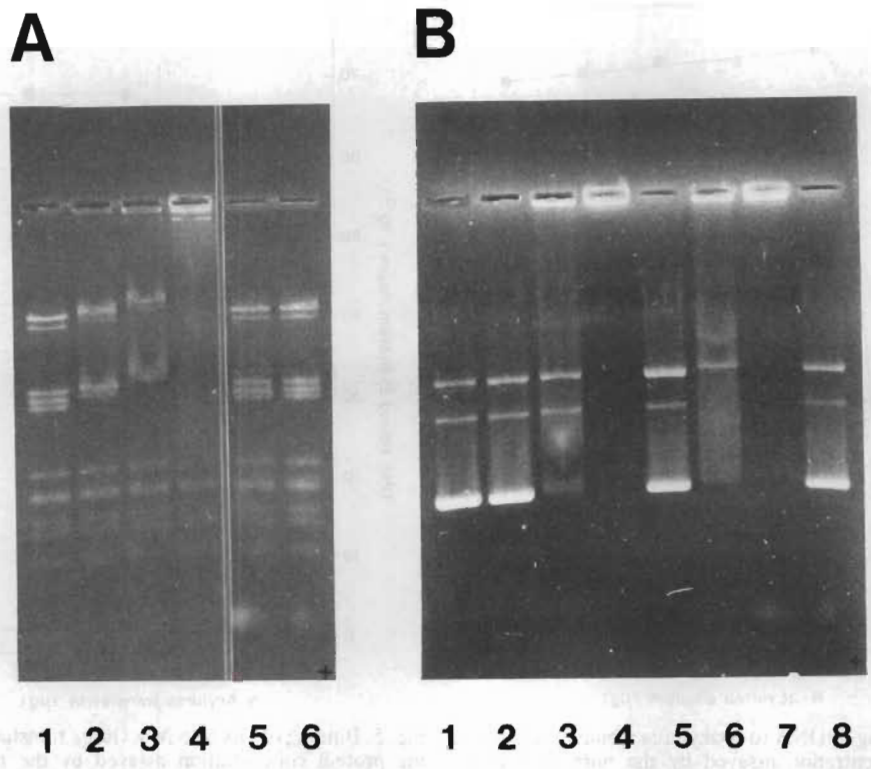


Fig. 3. Agarose gel electrophoresis of *N*-acylurea protein–DNA complexes. (*N*-Acylurea proteins prepared at a carbodiimide to protein mole ratio of 500:1.) (A) Interaction of *N*-acylurea CDI-transferrin with λ-PstI DNA restriction fragments. 1, λ-PstI DNA plus transferrin (4 μg); 2–4, λ-PstI DNA plus *N*-acylurea CDI-transferrin (2, 4 and 8 μg respectively); 5 and 6, λ-PstI DNA plus cytochrome C (4 and 8 μg respectively). λ-PstI DNA, 0.48 μg per well. (B) Interaction of *N*-acylurea CDI-proteins with pBR322 DNA. 1, pBR322 marker; 2, pBR322 plus albumin (4 μg); 3 and 4, pBR322 plus *N*-acylurea CDI-albumin (4 and 8 μg respectively); 5, pBR322 plus transferrin (4 μg); 6 and 7, pBR322 plus *N*-acylurea CDI-transferrin (4 and 8 μg respectively); 8, pBR322 plus cytochrome C (4 μg). pBR322 DNA, 0.33 μg per well.

similar binding characteristics to iron-free *N*-acylurea Me^{+}CDI-transferrin (Fig. 5).

Studies on the rate of binding of DNA to *N*-acylurea proteins was carried out at 5° and 20° using the nitrocellulose filter method of assay. At both temperatures the reaction was very fast, with binding occurring in less than 1 sec (results not given). However, as is shown later, this initial rapid binding appears to be weak, changing to a tight or more stable form of binding with time.

N-Acylurea proteins used in the binding experiments described above were all prepared at a carbodiimide to protein mole ratio of 500:1. A series of *N*-acylurea CDI-albumins varying in the carbodiimide to protein mole ratio employed during synthesis from 10:1 to 2000:1 was evaluated for binding ability by nitrocellulose filter assay. The binding of pBR322 DNA and calf thymus DNA by the various *N*-acylurea CDI-albumins is shown in Fig. 6. It should be noted that these particular binding assays were carried out in a buffer of higher ionic strength than that quoted previously; the buffer contained 0.05 M Tris–HCl (pH 7.5), 0.1 M NaCl and 5 mM EDTA. Results demonstrate that *N*-acylurea albumins prepared at CDI to protein mole ratios of 250:1 and greater are able to bind DNA.

The effect of NaCl *concentration on the binding of* N-*acylurea proteins to DNA*

A strong salt dependency of the binding between different types of protein and DNA has been observed [18, 23–26] and is indicative of electrostatic interaction between basic amino acids of the protein and phosphate groups of the nucleic acid. Structurally, the *N*-acylurea proteins (Fig. 2) are modified proteins with positively charged tertiary amino or quaternary ammonium groups attached via a spacer to glutamic and aspartic acid side chains of the protein. Electrostatic interaction between these groupings and DNA phosphates appeared likely and for this reason studies were carried out on the effect of salt concentration on DNA binding.

N-Acylurea proteins and DNA were incubated at 20° in a buffer consisting of 0.01 M Tris–HCl (pH 7.5), 0.05 M NaCl and 5 mM EDTA for 30 min; the solutions were then brought to the desired NaCl concentration, left for 5 min and finally filtered through nitrocellulose filters. Results of the salt-induced dissociations are shown in Fig. 7. Concentrations of NaCl required for the half-dissociation of individual *N*-acylurea protein–DNA complexes as deduced from the curves shown are: *N*-acylurea CDI-

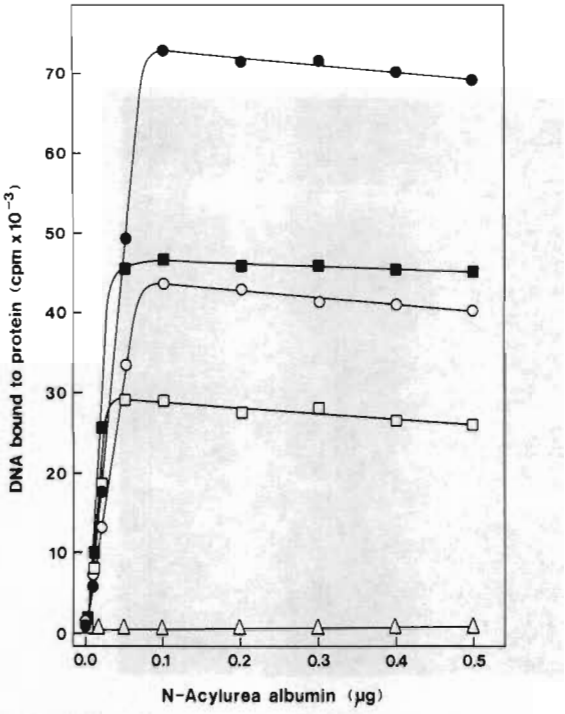


Fig. 4. Binding of DNA to *N*-acylurea albumins at varying protein concentration assayed by the nitrocellulose filtration method. DNA per reaction mixture, 0.01 μg. Modification of albumin was at a carbodiimide to protein mole ratio of 500:1. *N*-Acylurea CDI-albumin with [^{3}H]pBR322 DNA (●) and with [^{3}H]calf thymus DNA (○); *N*-acylurea Me$^+$CDI-albumin with [^{3}H]pBR322 DNA (■) and with [^{3}H]calf thymus DNA (□); albumin control with either DNA (△).

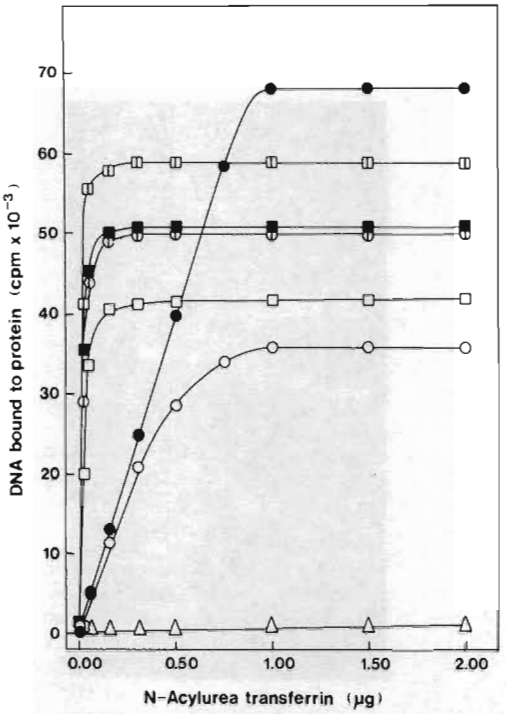


Fig. 5. Binding of DNA to *N*-acylurea transferrins at varying protein concentration assayed by the nitrocellulose filtration method. DNA per reaction mixture, 0.01 μg. Modification of the transferrin was at a carbodiimide to protein mole ratio of 500:1. *N*-Acylurea CDI-transferrin with [^{3}H]pBR322 DNA (●) and with [^{3}H]calf thymus DNA (○); *N*-acylurea Me$^+$CDI-transferrin with [^{3}H]pBR322 DNA (■) and with [^{3}H]calf thymus DNA (□); *N*-acylurea Me$^+$CDI-(Fe^{3+})transferrin with [^{3}H]pBR322 DNA (◫) and with [^{3}H]calf thymus DNA (◫); transferrin control with either DNA (△). The *N*-acylurea Me$^+$CDI-(Fe^{3+})transferrin used in these experiments was iron-loaded prior to carbodiimide modification.

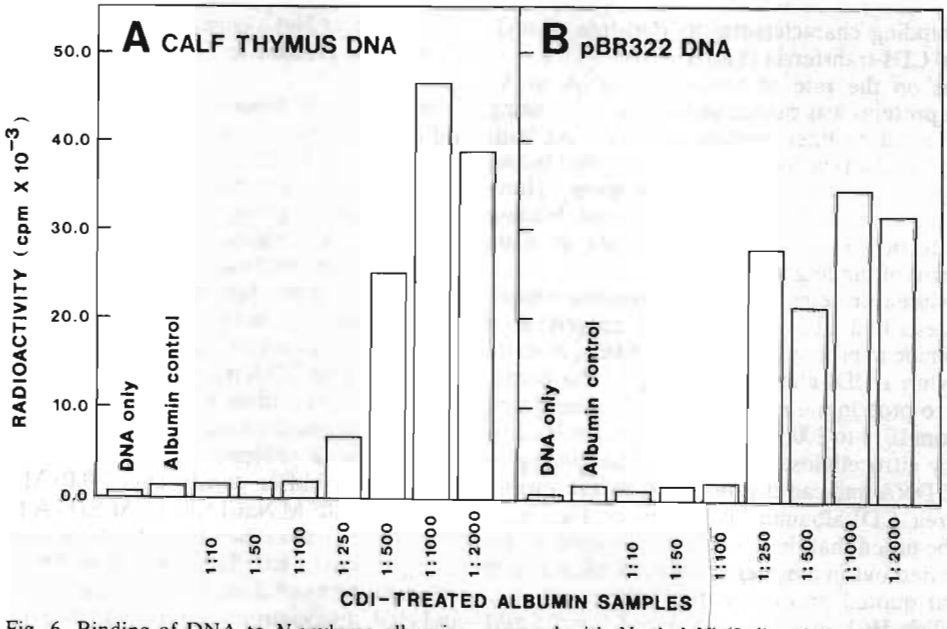


Fig. 6. Binding of DNA to *N*-acylurea albumins prepared with *N*-ethyl-*N'*-(3-dimethylaminopropyl)-carbodiimide at carbodiimide to protein mole ratios varying from 10:1 to 2000:1. All reaction mixtures contained 0.01 μg DNA and 8.8 μg protein. (A) [^{3}H]calf thymus DNA with the complete range of *N*-acylurea CDI-albumins. (B) [^{3}H]pBR322 DNA with the complete range of *N*-acylurea CDI-albumins.

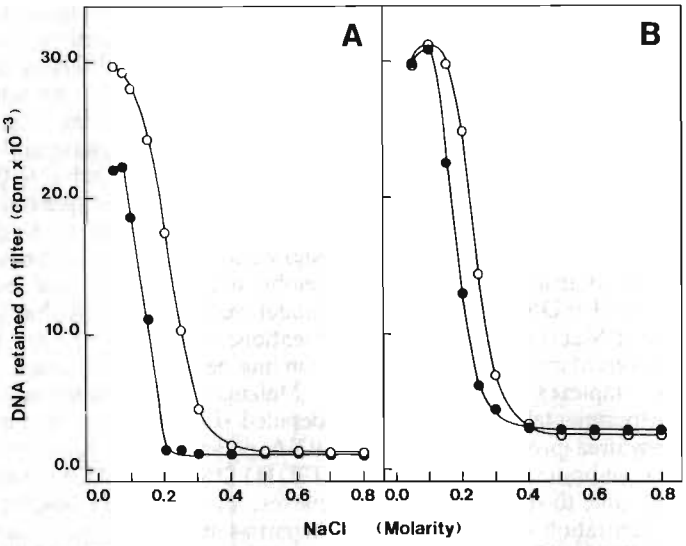


Fig. 7. The effect of NaCl concentration on the dissociation of *N*-acylurea protein–DNA complexes. *N*-Acylurea proteins were all prepared at a carbodiimide to protein mole ratio of 500:1. Protein per reaction mixture: *N*-Acylurea CDI-albumin, 0.10 μg; *N*-acylurea Me⁺CDI-albumin, 0.08 μg; *N*-acylurea CDI-transferrin, 2.0 μg; *N*-acylurea Me⁺CDI-transferrin, 0.20 μg. DNA per reaction mixture, 0.005 μg. (A) Dissociation of [^{3}H]pBR322 DNA from *N*-acylurea CDI-albumin (●) and *N*-acylurea Me⁺CDI-albumin (○). (B) Dissociation of [^{3}H]pBR322 DNA from *N*-acylurea CDI-transferrin (●) and *N*-acylurea Me⁺CDI-transferrin (○).

albumin, 0.15 M; *N*-acylurea Me⁺CDI-albumin, 0.24 M; *N*-acylurea CDI-transferrin, 0.18 M and *N*-acylurea Me⁺-CDI-transferrin, 0.26 M.

Heparin challenge experiments

Nitrocellulose filter retainable complexes are formed rapidly between *N*-acylurea proteins and DNA at both 5° and 20°. These initial complexes appear to be bound weakly, as they are dissociated if challenged by the polyanion heparin within 30 sec of formation (Fig. 8). However, formation of tightly bound complexes occurs if the binding reaction is allowed to proceed for 30 min at 20° or 37°. A heparin challenge at this time is significantly less effective in causing dissociation of the complex (Fig. 8).

Further investigations into the nature of binding

Further experiments aimed at obtaining information on the nature of the binding interactions between *N*-acylurea proteins and DNA were undertaken. Heating of *N*-acylurea CDI-albumin and *N*-acylurea CDI-transferrin at 55° for 10 min had no effect on their interaction with DNA. We were not able to extract pBR322 DNA from its complex with *N*-acylurea CDI-albumin by use of phenol in the absence of SDS. However, in the presence of 0.1% SDS at 60° for 5 min, pBR322 DNA was extracted into the aqueous phase during treatment with phenol, and recovered mainly in the supercoiled form (as shown by agarose gel electrophoresis). The results suggest that this particular *N*-acylurea protein–DNA complex is not held together by covalent bonds and that interaction does not involve the introduction of nicks into the DNA.

Competition experiments with various synthetic RNA polynucleotides produced inhibition of binding of pBR322 DNA to *N*-acylurea CDI-transferrin to differing degrees. Poly(G) produced the strongest inhibition, followed by poly(A), poly(U) and poly(C). Poly(m^{7}G) and poly(I) also inhibited binding but to a lesser degree than poly(G). The mononucleotides dGMP and dAMP were without effect on binding.

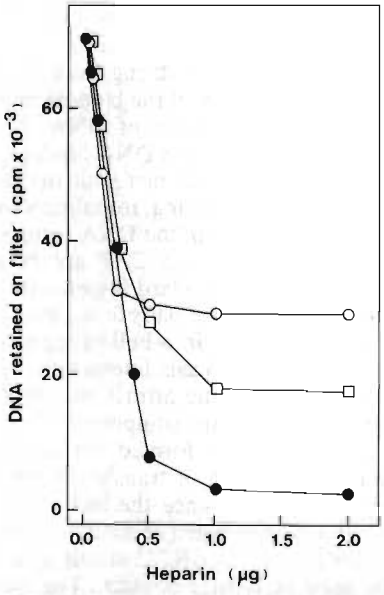


Fig. 8. The effect of heparin on the binding of DNA to *N*-acylurea transferrin. [^{3}H]pBR322 DNA (0.01 μg) and *N*-acylurea CDI-transferrin (500:1 mole ratio preparation; 4.0 μg) were incubated together for different times at either 20° or 37°. Each sample was then challenged with heparin for 30 sec and immediately processed by nitrocellulose filtration. Incubation for 0.5 min at 20° (●); incubation for 30 min at 20° (□); incubation for 30 min at 37° (○).

DISCUSSION

The work presented in this paper describes the modification of albumin and transferrin with water-soluble carbodiimides to give *N*-acylurea proteins which are able to bind various types of DNA, including λ-PstI restriction fragments, pBR322 plasmid, calf thymus DNA and single-stranded M13 mp8 DNA.

Binding reactions were investigated by two methods. Initial experiments made use of agarose gel electrophoresis to show that DNA migration was retarded in the presence of *N*-acylurea proteins, and that at higher concentrations of modified protein, *N*-acylurea protein–DNA complexes were not able to enter the gel. Further experimental work was based on the binding of *N*-acylurea protein–DNA complexes to nitrocellulose membrane filters. Using this type of assay we were (i) able to study the effect of *N*-acylurea protein concentration on its binding to calf thymus and pBR322 DNA, (ii) measure the rate of formation of complexes, which is rapid, (iii) show that *N*-acylurea protein–DNA interactions are sensitive to salt concentration, and (iv) distinguish two modes of interaction between *N*-acylurea proteins and DNA.

Additional experiments involving phenol extraction of *N*-acylurea protein–DNA complexes in the presence of SDS showed that pBR322 could be recovered in an intact supercoiled form. This information, together with the results of salt dissociations and heparin competition experiments, suggests that binding is electrostatic in nature, although other forms of binding might also be involved.

Results presented in Fig. 7 highlight the dependence of the apparent association constant (K_{obs}) for the *N*-acylurea albumin-pBR322 and *N*-acylurea transferrin-pBR322 interactions on monovalent ion concentration [M^+]. The reduction of K_{obs} with increasing [M^+] suggests a strong electrostatic component in the free energy of the binding interaction.

Assuming B-conformation of DNA in solution [25], the geometry of known DNA binding proteins allows for considerable contact with the duplex as determined by model building, revealing a large non-electrostatic component in the DNA complex. Thus Cro, lambda repressor and CAP are believed to interact with DNA of standard B-geometry through hydrogen bonding of base pairs in the major groove with specific side chains in α-helical regions located therein [26]. Van der Waals interactions as well as the expected electrostatic attractions are apparent [26]. The non-electrostatic component of the ligand–nucleic acid complexes formed between carbodiimide-modified albumin or transferrin and DNA is believed to be minor, since the lack of detectable complex formation between unmodified albumin and transferrin with either pBR322 or calf thymus DNA indicates poor molecular contact. The concept of relative surface exposure of nucleic acids, which has not been developed as extensively as with proteins, has been investigated by Alden and Kim [27] to identify the most accessible atoms or regions of nucleic acids available for intermolecular interactions. These studies reveal that the accessible area in B-DNA is rather polar, with phosphate oxygens accounting for 45% of this surface. Their peripheral location further supports the notion that the charge–charge interactions observed in our studies between carbodiimide-modified proteins and DNA are accounted for in large measure by ion pairs formed between DNA phosphate and the nitrogens of the *N*-acylurea groups attached to the proteins. This may not be the case with single-stranded M13 mp8 since the accessible surface of single-stranded DNA is known to be larger and considerably more hydrophobic due to greater base exposure [27]. Indeed, model building suggests that ion pairs of the type mentioned may be stabilized by specific hydrogen bonding between *N*-acylureas and the bases.

Melancon and co-workers [23] have carried out detailed studies on the binding of *Escherichia coli* RNA polymerase to restriction enzyme digests of T7D111 DNA. In addition to normal promoter complexes, these authors observed both fast-forming heparin-sensitive complexes and polymerase–DNA complexes insensitive to challenge with heparin but formed at a slower rate. It is of interest to note that in this respect our synthetic *N*-acylurea proteins form complexes with DNA which have similarities to the non-promoter complexes found in the bacterial system (Fig. 8).

The preparation of nucleic acid–protein conjugates for use as immunogens has been outlined by Stollar [28]. In one method, methylated serum albumin which is positively charged is used to prepare insoluble complexes with DNA. In an alternative method, a nucleotide or oligonucleotide is reacted with a water-soluble carbodiimide and then added to a protein for coupling; the nucleotide or oligonucleotide appears to become linked to the protein through a covalent phosphoramidate bond. Interestingly, Stollar notes that in certain cases an excess of carbodiimide modifies the protein carboxyl groups such that the resultant protein forms a precipitate with DNA.

Cheng *et al.* [7] have recently suggested that the introduction of DNA into cells with genetic defects may be achieved by linking the DNA to a protein which can be bound by specific cell surface receptors. Internalization via receptor-mediated endocytosis might thus allow entry of the DNA into a specific cell type. These workers have devised a method which involves chemical modification of both protein and DNA components, followed by the covalent linking of the two moieties through disulphide exchange reactions. The finding that *N*-acylurea albumin was able to bind DNA non-covalently and reversibly suggested that carbodiimide-modification of transferrin and other polypeptides which react with cell surface receptors might give *N*-acylurea substituted molecules that are able to bind DNA and still interact successfully with their receptors, thereby affording an alternative approach to this possible method of DNA (gene) transfer [7, 8]. Reaction of iron-free and iron-loaded transferrin with carbodiimides was therefore undertaken. It is notable that the resultant *N*-acylurea transferrins, both iron-free and iron-loaded, bind DNA to a similar degree (Fig. 5).

Work on the interaction of *N*-acylurea protein–DNA complexes with cells is in progress.

Acknowledgement—The authors thank Professor D. Grunberger of Columbia University, New York, for his encouragement in this work.

REFERENCES

1. T. Okabayashi, S. Mihara, D. B. Repke and J. G. Moffatt, *Cancer Res.* **37**, 619 (1977).
2. T. Okabayashi, S. Mihara and J. G. Moffatt, *Cancer Res.* **37**, 625 (1977).
3. S. Bauminger, S. Wilchek and M. Wilchek, in *Methods in Enzymology*, Vol. 70 (Eds. H. Van Vunakis and J. J. Langone), p. 151. Academic Press, New York (1980).
4. F. Millet, V. Darley-Usmar and R. A. Capaldi, *Biochemistry* **21**, 3857 (1982).
5. L. M. Geren, P. O'Brien, J. Stonehuerner and F. Millet, *J. biol. Chem.* **259**, 2155 (1984).
6. K. L. Carraway and D. E. Koshland, in *Methods in Enzymology*, Vol. 25 (Eds. C. H. W. Hirs and S. N. Timasheff), p. 616. Academic Press, New York (1972).
7. S. Y. Cheng, G. T. Merlino and I. Pastan, *Nucleic Acids Res.* **11**, 659 (1983).
8. B. C. F. Chu, G. M. Wahl and L. E. Orgel, *Nucleic Acids Res.* **11**, 6513 (1983).
9. P. W. J. Rigby, M. Dreckmann, C. Rhodes and P. Berg, *J. molec. Biol.* **113**, 237 (1977).
10. T. Maniatis, E. F. Fritsch and J. Sambrook, in *Molecular Cloning*, p. 466. Cold Spring Harbour Laboratory (1982).
11. O. H. Lowry, N. J. Rosebrough, A. L. Farr and R. J. Randall, *J. biol. Chem.* **193**, 265 (1951).
12. M. G. Kopczynski and B. M. Babior, *J. biol. Chem.* **259**, 7652 (1984).
13. J. H. Ward, J. P. Kushner and J. Kaplan, *J. biol. Chem.* **257**, 10317 (1982).
14. R. F. Nunn, in *Scientific Foundations of Clinical Biochemistry* (Eds. D. L. Williams, R. F. Nunn and V. Marks), p. 33. William Heinemann, London (1978).
15. U. K. Laemmli, *Nature* **227**, 680 (1970).
16. R. Timkovich, *Analyt. Biochem.* **79**, 135 (1977).
17. H. A. Dailey and P. Strittmatter, *J. biol. Chem.* **254**, 5388 (1979).
18. D. C. Hinkle and M. J. Chamberlin, *J. molec. Biol.* **70**, 157 (1972).
19. A. D. Riggs, H. Suzuki and S. Bourgeois, *J. molec. Biol.* **48**, 67 (1970).
20. A. D. Johnson, C. O. Pabo and R. T. Sauer, in *Methods in Enzymology*, Vol. 65 (Eds. L. Grossman and K. Moldave), p. 839. Academic Press, New York (1980).
21. G. Zubay, in *Methods in Enzymology*, Vol. 65 (Eds. L. Grossman and K. Moldave), p. 856. Academic Press, New York (1980).
22. R. M. Myers and R. Tjian, *Proc. natn Acad. Sci. U.S.A.* **77**, 6491 (1980).
23. P. Melancon, R. R. Burgess and M. T. Record, *Biochemistry* **21**, 4318 (1982).
24. R. L. Karpel, D. J. Merkler, B. K. Flowers and M. D. Delahunty, *Biochim. biophys. Acta* **654**, 42 (1981).
25. J. Wang, *Proc. natn Acad. Sci. U.S.A.* **76**, 200 (1979).
26. C. O. Pabo and R. T. Sauer, in *Annual Review of Biochemistry*, Vol. 53 (Eds. C. C. Richardson, P. D. Boyer and A. Meister), p. 293. Annual Reviews Inc., Palo Alto, California (1984).
27. C. J. Alden and S-H. Kim, in *Stereodynamics of Molecular Systems* (Ed. R. H. Sarma), p. 331. Pergamon Press, Oxford (1979).
28. B. D. Stollar, in *Methods in Enzymology*, Vol. 70 (Eds. H. Van Vunakis and J. J. Langone), p. 70. Academic Press, New York (1980).

EVIDENCE FOR TARGETED GENE TRANSFER BY RECEPTOR-MEDIATED ENDOCYTOSIS: STABLE EXPRESSION FOLLOWING INSULIN-DIRECTED ENTRY OF *NEO* INTO HepG2 CELLS

Running Title: GENE TRANSFER BY RECEPTOR-MEDIATED ENDOCYTOSIS

Barbara Huckett, Mario Ariatti and Arthur O. Hawtrey#

Department of Biochemistry, University of Durban-Westville,
Private Bag X54001, Durban 4000, South Africa.

Present address: Department of Pharmacology,
Faculty of Medicine,
University of Stellenbosch,
P. O. Box 63,
Tygerberg 7505,
South Africa.

ABSTRACT

Evidence is presented for targeted gene delivery to HepG2 cells via the endocytotic pathway under the direction of insulin. Serum albumin was treated with the water-soluble carbodiimide *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride and the resultant positively charged *N*-acylurea albumin covalently conjugated to insulin by glutaraldehyde cross-linkage. The conjugated protein is shown by nitrocellulose filter binding and gel band shift assays to bind DNA, and by competitive displacement of [^{125}I]insulin to bind to the insulin receptor. When the expression vectors ptkNEO and pAL-8 which incorporate the *neo* gene were complexed to the conjugate in an *in vitro* system of transfection, G418 resistant clones developed at frequencies of 2.0 - 5.5 X 10^{-5}. Subsequently, a 923bp PstI fragment within the *neo* sequence was identified by Southern transfer in genomic DNA from transfected cell populations which had been maintained on a G418 regime for 44 days.

Abbreviations: BSA, bovine serum albumin; CDI, *N*-ethyl-*N'*-(3-dimethylaminopropyl)-carbodiimide; DMSO, dimethyl sulphoxide; HEPES, 2-[4-(2-hydroxyethyl)-1-piperazinyl]-ethanesulphonic acid; SSC, 0.15M NaCl, 0.015M trisodium citrate.

INTRODUCTION

The deliberate insertion and integration of functionally active genes into mammalian cells is now a commonplace procedure which has manifold application in molecular biology and medicine. The transfer of specific genes permits: study of the functional domains of proteins [1], fine dissection of the mechanisms by which gene expression is regulated [2], elucidation of cellular differentiation hierarchies by the introduction of clonal markers into embryonic cells or primitive stem cells [3], analysis of oncogene and putative oncogene function in carcinogenesis [4], the production of transgenic animals [5] and the correction of genetic disorders [6].

A variety of methodologies exist for the necessary transfer of exogenous DNA to the cell interior. The most straightforward and commonly used are chemical methods such as those involving treatment with calcium phosphate [7] or a cationic facilitator [8] which induce indiscriminate DNA attachment to the membrane and thereby give rise to endocytotic uptake. It has been shown recently, however, that the calcium phosphate transfection protocol has been implicated in perturbations of gene expression [9, 10]. Other popular methods of gene transfer involve active injection of DNA during physical puncture [11] or passive uptake during poration or abrasion of the membrane in the presence of DNA [12, 13] but, while moderately efficient, these are open to

some criticism on the grounds that they are intrinsically aggressive. Also, as with the chemical techniques, they are applicable only to *in vitro* cell systems. A further category of techniques relies upon the mediation of DNA transfer via membrane-bound vesicles such as intact protoplasts [14], erythrocyte ghosts [15], reconstituted viral envelopes [16] and liposomes [17], the last of these having the design capacity to be adapted for targeting to specific cell types both *in vitro* and *in vivo* [18]. In a class of their own are viral vectors: reconstructed viruses, especially retroviruses, disarmed and expropriated for the carriage, insertion and precisely controlled genomic integration of DNA sequences [19]. Such vectors constitute the most efficient means of gene transfer known at the present time, but the technicalities are too intricate for many straightforward applications and there are numerous difficulties still existing which must be resolved before this approach can be widely advocated [20].

In this communication we present a procedure for gene transfer which exploits the efficiency and specificity of internalisation afforded by the process of receptor-mediated endocytosis [21]. DNA in the naked state is carried into the cell reversibly bound to a soluble protein ligand recognisable by its cognate receptor at the plasma membrane surface. The non-covalent complexing of DNA to ligand is achieved by modifying the protein with a water-soluble carbodiimide, *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (CDI) under controlled conditions which allow the formation of basic *N*-acylurea moieties at the carboxyl groups of available aspartate and glutamate residues [22]. The resultant electropositive *N*-acylurea protein readily interacts with ionised phosphodiester backbone regions of DNA and tenaceous salt bridges are formed [23]. In the case of small polypeptide ligands, modification of the ligand itself is minimised by covalently attaching it in the unmodified state to *N*-acylurea serum albumin, thereby creating a conjugate which relegates the receptor binding and DNA binding functions to separate protein surfaces on one macromolecule (Figure 1).

Since the modified ligands are extensively purified prior to use in transfection, and are applied together with unmodified DNA to recipient cells at low concentration in medium which requires no special adjuvants, they mimic the behaviour of their native counterparts and undergo natural endocytotic uptake in a highly directed and non-inimical manner. Prudent choice of ligand could possibly make this method of gene transfer applicable to a variety of cell lines in culture. Furthermore, the potential exists to develop the concept for targeting genes to specific tissue types in whole animal investigations.

In this study we have modelled the approach using insulin as the ligand of choice. We describe the preparation of insulin-[*N*-acylurea albumin] conjugates, their binding to insulin receptors on cultured HepG2 cells, and their application in the uptake into and expression in HepG2 cells of the *neo* gene from the bacterial transposon Tn5 incorporated into the expression vectors ptkNEO and pAL-8.

MATERIALS AND METHODS

Materials. Bacteriological culture media and ingredients were obtained from Difco Labs, cell culture media from GibCo, and sterile plasticware from CelCult. Filter sterilisation units and nitrocellulose discs were from Millipore. Monocomponent porcine insulin was purchased from Novo Biolabs; bovine serum albumin (fraction V), DNA polymerase I (EC 2.7.7.7), DNA polymerase I Klenow fragment, restriction endonucleases and plasmid pBR322 from Boehringer-Mannheim; DNase I (EC 3.1.21.1), lysozyme (EC 3.2.1.17), ethidium bromide, ampicillin, chloramphenicol and Geneticin (G418) from the Sigma Chemical Co. [Methyl,1′,2′-^{3}H] thymidine 5′-triphosphate (specific radioactivity 93 Ci/mmol), thymidine 5′-[α-^{32}P] triphosphate (specific radioactivity 3000 Ci/mmol), [^{125}I-tyrosine A14] human insulin (specific radioactivity 2000 Ci/mmol), Hybond C-extra and Hyper-film MP were obtained from Amersham, U.K. Anti-porcine insulin was purchased from Bio-Yeda. *N*-Ethyl-*N′*-(3-dimethylaminopropyl) carbodiimide hydrochloride and caesium chloride were from Merck. Agarose (ultra pure grade and low melting temperature grade), acrylamide, bis-acrylamide, nitrocellulose sheets and Immun-Blot kit were obtained from BioRad. Sephadex G-50 and G-100 were from Pharmacia. All other reagents were of analytical grade.

Maintenance of HepG2 cells. HepG2 cells [24] were grown at 37°C in flasks containing Eagle's minimal essential medium supplemented with 10mM $NaHCO_3$ and buffered with 20mM Na-HEPES, pH 7.3 (MEM) containing 10%v/v foetal calf serum (MEM+S). Cells at semi-confluence (4-7 days' growth) were trypsinised with a solution containing 0.25%w/v trypsin, 0.1%w/v EDTA, and the suspensions subcultured at 1/6 or 1/3 dilutions.

Generation of expression vectors. The expression vector plasmids ptkNEO and pAL-8 were each propagated in *Escherichia coli* strain HB101. Initial transformation was achieved using a simple calcium method [25] and stocks of the respectively transformed host strain kept on ampicillin media as both anaerobic stab cultures and frozen glycerol-broth cultures. Large scale plasmid production followed conventional protocols [26], extraction of plasmids was by the boiling method [26] or the lysozyme-Triton X-100 method of Katz *et al* [27] and purification was by centrifugation in caesium chloride-ethidium bromide (CsCl-EB) density gradients [28]. After EB removal from the plasmid fractions and subsequent dialysis against TE buffer (10mM Tris-HCl, 1mM EDTA, pH 7.5) the final preparations were stored over chloroform (5%v/v) at -15°C.

Quantification of DNA. DNA was estimated spectrophotometrically by UV absorption.

Agarose gel electrophoresis. Plasmids and restriction digests of plasmids were analysed on 1.2% agarose gels run at pH 7.5 and stained with EB using conventional methodology [26]. Preparative electrophoresis for the separation and excision of restriction fragments for subsequent labelling as probes was

performed in a similar manner on gels of low melting temperature agarose. For the analysis of DNA-protein binding by band shift assay, plasmids were bound to proteins prior to electrophoresis as described in the legend to Figure 5. For the analysis of the charge properties of proteins, protein samples were electrophoresed as above but subsequently fixed in 10%w/v TCA, stained with 0.025%w/v aqueous Coomassie Brilliant Blue in 10%w/v TCA, and destained in 10%w/v TCA, 5%v/v acetic acid.

Labelling of pBR322. The plasmid pBR322 was labelled by the nick translation protocol [29] in the presence of [methyl, 1′,2′-^{3}H] thymidine 5′-triphosphate and subsequently purified by the spun column method using Sephadex G-50 [26]. The product was stored at -15°C. Specific activity was 6.0 - 8.5 X 10^{6} cpm/µg TCA-precipitated DNA.

Labelling of neo probe. A 923bp PstI fragment of ptkNEO representing the majority of the *neo* gene sequence and no extraneous sequences was labelled with thymidine 5′-[α-^{32}P] triphosphate by the method of Feinberg and Vogelstein [30]. The desired fragment was excised from an electrophoretic gel of low melting temperature agarose, water added (3µl/mg gel), and the sample boiled for 7 minutes immediately prior to use in the labelling reaction. The reaction product was subjected to spun column purification using Sephadex G-50 (25) then stored at -15°C. Specific activity was 6.0 X 10^{8} - 2.2 X 10^{9} cpm/µg TCA-precipitated DNA.

Preparation of N-acylurea albumins. The *N*-acylurea derivative of bovine serum albumin (BSA) was prepared as previously described [23] by reaction of the protein with *N*-ethyl-*N*′-(3-dimethylaminopropyl) carbodiimide HCl (CDI), at a 1:500 mole ratio. The product was dialysed exhaustively against saline-EDTA (50mM NaCl, 0.3mM EDTA, pH 7.0) at 5°C and aliquots stored at -15°C.

Conjugation of insulin to albumins. Insulin was cross-linked to BSA at a 10:1 mole ratio by a development of the method of Poznansky *et al* [31]. Insulin (4.2mg, 7.37 X 10^{-7} moles, dissolved in 600µl DMSO at 37°C) was mixed with unmodified BSA or its *N*-acylurea derivative (5mg, 7.35 X 10^{-8} moles, in 1ml saline-EDTA) at 20°C then diluted with 4.0ml 50mM acetate buffer, pH 3.6, and the solution cooled to 5°C. Glutaraldehyde (90µl, 25%) was added and the reaction mixture maintained in the cold overnight. Thereafter, crystalline glycine (75mg) was added, dissolved, and allowed to react with excess glutaraldehyde for 2.5h at 20°C. The reaction mixture was then dialysed exhaustively against 10mM acetate buffer, pH 4.0, containing 5%v/v DMSO at 5°C. The resultant clear, pale straw-coloured solutions were stored in aliquots at -15°C.

Fractionation of conjugates. Conjugates were purified by gel filtration through Sephadex G-100. Fractions (1ml) were stored at -15°C for up to 8 weeks. Conjugates for use in transfection experiments were fractionated under sterile conditions.

Quantification of protein. Protein was estimated by the method of Lowry *et al* [32] using BSA as standard.

Identification of insulin in the conjugates. Insulin was detected semi-quantitatively in the conjugate fractions by immuno-dot-blotting. Each fraction tested was diluted in elution buffer to a concentration of 66ng protein/µl. These starting solutions, together with a range of double dilutions derived from them, were applied in 1µl aliquots to nitrocellulose and processed according to the BioRad Immun-Blot procedure using anti-insulin (1/333) as first binding agent, protein A-horseradish peroxidase (1/2500) as second binding agent, and a mixture of 4-chloro-1-naphthol (dissolved in methanol) and hydrogen peroxide (aqueous) as colour reagent.

Molecular size determination. Protein conjugates were analysed by SDS-polyacrylamide gel electrophoresis according to the method of Laemmli [33].

DNA binding assay. The binding of DNA to protein conjugate fractions was estimated by filtration through nitrocellulose in a variation of the method previously described [23]. Mixtures containing 50mM NaCl, 20mM Tris-HCl (pH 7.0), 1mM EDTA, [^{3}H]pBR322 (4.5ng, 3.3 X 10^{4} cpm), unlabelled pBR322 (5.5ng) and protein as indicated in the figure legends in a final volume of 200µl were incubated at 20°C for 30min, then filtered through nitrocellulose discs (Millipore Type HA, 25mm diameter, 0.45µm pore size) and rinsed with 2.0ml of buffer consisting of 50mM NaCl, 20mM Tris-HCl (pH 7.0), 1mM EDTA. The nitrocellulose filters were wetted in rinse buffer prior to use. Finally, filters were air-dried and counted for radioactivity in scintillation fluid.

Receptor binding studies. HepG2 cells were heavily seeded (1/3 splits) into 35mm diameter well plates and grown to semi-confluence. Cells were washed twice with MEM (no serum) at 37°C (2ml per well) then incubated at 37°C for 1h in the presence of a further 2ml MEM, which was subsequently removed. Plates were placed on ice and MEM at 5°C added (1ml per well) together with [^{125}I]insulin and unlabelled insulin or insulin conjugate as indicated in the figure legends. Additives were thoroughly mixed into the binding medium and plates incubated at 10°C. Cells were subsequently washed three times with cold phosphate buffered saline (PBS) (4ml per well) and drained. After addition of water (1ml per well), cells were loosened into suspension with a rubber policeman, 200µl removed for protein determination, and the remainder fully dispersed by the addition of 1ml lysis buffer (0.5% SDS, 100mM NaCl, 40mM Tris-HCl, 20mM EDTA, pH 7.0) [34]. The lysates were transferred to tubes for gamma counting.

Transfection procedure. HepG2 cells were heavily seeded (1/3 splits) in 25cm^{2} flasks and grown to semi-confluence. Cells were washed three times with 4ml MEM (no serum), then incubated at 37°C for 1h in the presence of a further 4ml MEM, which was subsequently removed. Flasks were placed on ice and 2ml MEM at

5°C added together with I-[A-CDI] or control protein as indicated in the legend to Figure 8. After overnight incubation at 10°C, flasks were again placed on ice, the medium removed, and 1ml MEM at 5°C added together with expression vector as indicated. The vector samples used comprised superhelical and nicked circular forms of plasmid DNA in approximately equal proportions. Flasks were transferred from ice to incubation at 10°C (30min), 20°C (30min), then 37°C (5.0h). At this stage 3ml MEM+S (plus serum) was added and incubation continued for 24h. Flasks were trypsinised and duplicate subcultures established containing 10^6 cells per flask. Growth was continued for a further 24h before the first addition of G418.

Stable expression of NEO: G418 selection. Following the trypsinisation of post-transfection cells, G418 stock solution (24mg total antibiotic/ml H_2O adjusted to pH 7.4 with NaOH and filter sterilised) was added to the medium to give a final concentration of 1200µg/ml (biological activity = 50% of this) and incubation continued for 5 days, after which the medium was changed and fresh G418 added at a final concentration of 600µg/ml. Medium changes were effected at approximately 5 day intervals thereafter, the G418 level being lowered further to and maintained at 500µg/ml. Following the unambiguous appearance of resistant clones (14-21 days) and the death of most non-resistant cells (21-28 days), one set of flasks was stained for the recording of clone distribution. Surviving cells in the other set were trypsinised for subculture and propagation in a G418 maintenance regime and ultimately subjected to genomic
DNA extraction.

Extraction of genomic DNA from HepG2 cells. Duplicate 75cm^2 flasks of cells were used for each DNA preparation. Extraction was according to the method of Shih and Weinberg [34]. Final DNA solutions were dialysed exhaustively against 0.1SSC at 5°C then stored at -15°C.

Detection of the integrated neo sequence in transfected HepG2 DNA. This was carried out by Southern hybridisation [35]. Blots were probed with a 923bp *neo* fragment and washed under conditions of high stringency. Bands were visualised by autoradiography.

RESULTS

Preparation of N-acylurea albumin

N-Acylurea albumin (A-CDI) was prepared by treating BSA with the water-soluble carbodiimide *N*-ethyl-*N'*-(3-dimethylaminopropyl) carbodiimide hydrochloride (CDI), under conditions previously described, in which 27 out of a total of 97 side chain carboxyl groups per BSA molecule become substituted by *N*-acylurea moieties leading to significantly increased electropositivity as shown by agarose gel electrophoresis at pH 7.5 [23]. During the reaction the carbodiimide induces a degree of polymerization of

albumin as shown by the behaviour of A-CDI on gel filtration (Figure 2.).

Preparation of insulin-albumin conjugates

Monocomponent porcine insulin (I) was covalently cross-linked to BSA (A) and N-acylurea BSA (A-CDI) by reaction with glutaraldehyde to give rise to the conjugates I-A and I-[A-CDI] respectively (Figure 1.). In order to ensure the attachment of several molecules of insulin (mol wt 5700) to each molecule of albumin (mol wt 68000) reaction mixtures contained an insulin to BSA mole ratio of 10:1 (mole calculations based on unmodified BSA input in both cases). Reaction and subsequent purification was conducted at low pH (3.6-4.0) in the presence of DMSO (5-10%v/v) in order to maintain insulin solubility at the required protein concentration. Gel filtration of the dialysed reaction mixtures through Sephadex G-100 showed the two reaction products to have identical profiles representing a series of high molecular weight complexes with only insignificant amounts of protein evident at the position of unconjugated insulin elution (Figure 2.). The incorporation of insulin into the conjugate fraction was confirmed by SDS polyacrylamide gel electrophoresis (results not shown). Furthermore, presence of insulin in the high molecular weight product was demonstrated unambiguously by immuno-dot-blot analysis, with the highest ratios of insulin to total protein being apparent in fractions 18 to 24 (Figure 2.). Results (not shown) of agarose gel electrophoresis at pH 7.5 of the conjugate fractions bore out the anticipated charge differences between I-A and I-[A-CDI], the latter showing considerably increased electropositivity due to the presence of basic *N*-acylurea moieties attached to surface aspartate and glutamate residues.

Binding of DNA to insulin-albumin conjugates

Purified conjugates obtained by gel filtration through Sephadex G-100 were assayed for DNA binding activity by the nitrocellulose filtration technique. The binding of DNA to *N*-acylurea proteins has been shown to be salt concentration dependent [23]; consequently the ionic strength of the incubation and washing medium must be carefully controlled. In addition, both protein-DNA binding and protein-nitrocellulose binding are pH dependent, and it was found to be important to conduct these assays between pH 6.0 and pH 7.5 (results not shown). Accordingly, for these experiments, an assay buffer was devised which maintained suitable ionic strength and pH despite the addition of variable volumes of conjugate dissolved in 10mM acetate buffer, pH 4.0. The assay buffer used was 50mM NaCl, 20mM Tris-HCl (pH 7.0), 1mM EDTA.

Figure 3. illustrates typical binding interactions of pBR322 DNA with the various size species of I-A and I-[A-CDI] obtained by gel filtration. Each assay mixture contained 1.5µg conjugate. At this level of protein the DNA present in the assay mixture (10ng) is not limiting (refer to Figure 4.). The capacity to bind DNA is demonstrated by I-[A-CDI] but not I-A. While all

fractions of I-[A-CDI] undergo significant binding, binding is lowest for the very large polymers and there is a reproducible trend for the binding capacity of the larger species of conjugate to be inversely proportional to molecular size (fractions 15 - 20). This is likely to be a reflection of the reduction in the ratio of surface area to weight that accompanies polymerization. Maximal binding is achieved and maintained over a range of the more moderately sized species (fractions 21 - 24).

On the basis of (i) DNA binding capacity, (ii) insulin content, and (iii) concentration, I-[A-CDI] fractions 19 to 23 inclusive were pooled to provide an I-[A-CDI] stock solution for use in all subsequent experiments. Fractions of I-A were pooled in an equivalent manner. Stock solution concentrations ranged from 0.38 to 0.42µg protein/µl.

The interaction of pBR322 DNA with insulin-albumin conjugates at varying concentrations of protein is shown in Figure 4. Results confirm that only I-[A-CDI] undergoes binding. Total DNA per reaction mixture is 10ng. The amount of DNA bound is seen to increase linearly as the input of I-[A-CDI] rises to 6µg; maximum binding is reached at 8µg I-[A-CDI].

DNA binding by I-[A-CDI] is illustrated by alternative means in Figure 5. which shows the result of DNA gel electrophoresis following incubation of pBR322 samples (0.45µg) with varying amounts of insulin-albumin conjugates. Retardation of migration resulting from the complexing of DNA to protein starts to become visible at 1.0µg I-[A-CDI], particularly in relation to the supercoil band, and at 1.66µg I-[A-CDI] significant retardation of both supercoil and nicked circular forms is seen as continuous smudging of the DNA. At higher levels of I-[A-CDI] a proportion of the complex is virtually unable to enter the gel. In contrast, I-A is shown to bring about no significant retardation of migration. Unmodified albumin and the basic protein cytochrome C are without effect on pBR322 migration under the same electrophoretic conditions [23]. Insulin is similarly without effect (result not shown).

Binding of insulin and insulin-albumin conjugates to the insulin receptor

Since laboratory iodination of insulin with [^{125}I] incurs oxidation damage and brings about indiscriminate labelling of tyrosine residues leading to decreased biological activity including aberrant receptor binding [36], commercially produced human insulin [^{125}I]-labelled only at the A14 tyrosine was employed in these investigations. Monoiodo[Tyr^{A14}]insulin is fully active in receptor binding assays [36]. The binding to insulin receptors of unlabelled porcine insulin and insulin-albumin conjugate derivatives was studied by determining their ability to compete with free radioactively labelled human insulin for receptor sites on HepG2 cells. Preliminary experiments were carried out, therefore, to establish the normal binding characteristics of [^{125}I-Tyr^{A14}] human insulin to HepG2

cells.

Prior to all binding procedures cells were washed to remove extracellular ligand, incubated at 37°C for 1h in serum-free medium in order to clear the receptors of bound ligand, then cooled. Temperatures below 16°C are known to inhibit absorptive endocytosis of insulin [37]. Accordingly, binding incubations were set up in serum-free medium at 10°C.

Figure 6A. presents the time course of binding of human insulin to HepG2 cells at 10°C. Binding takes place rapidly for the first 2h, after which it slows down, reaching a maximum at approximately 16h. For all subsequent binding experiments incubation was for an 18h period at 10°C. The effect of ligand concentration on the binding of human insulin to HepG2 cells is shown in Figure 6B. Binding increases with insulin input, and at an insulin concentration of 0.8ng/ml, saturation of the receptor population in the system is not yet reached.

Porcine insulin as well as insulin-albumin conjugates derived from porcine insulin compete effectively with human insulin for specific receptor sites on HepG2 cells (Figure 7.). Binding resulting from the addition of 0.4ng [^{125}I-Tyr^{A14}] human insulin alone is reduced to 50% in the presence of 50ng porcine insulin (Figure 7A.), 300ng I-A and 200ng I-[A-CDI] (Figure 7B.). By contrast, albumin unconjugated to insulin does not compete for binding. Furthermore, Figure 7C. shows that when I-[A-CDI] is prebound to DNA its receptor binding facility is not affected. DNA itself has no affinity for the binding site.

Insulin-[N-acylurea albumin] mediated transfection

DNA transfer to HepG2 cells was tested using the expression vectors ptkNEO and pAL-8, both of which carry the bacterial sequence *neo* from transposon Tn5 which encodes aminoglycoside 3'-phosphotransferase II. This gene is normally absent in mammalian cells but if inserted and expressed confers upon them resistance to the lethal effects of the 2-deoxystreptamine antibiotic G418 (Geneticin), thus acting as a dominant selective marker [38]. Viral promoter sequences allow expression of *neo* on transfer: in ptkNEO the gene is linked on either side to HSV thymidine kinase regulatory sequences; in pAL-8 the flanking transcriptional control regions are derived from SV40.

Figure 8. illustrates clonal colonies which have arisen over a period of 28 days in the presence of G418, indicating reception of the *neo* gene and its stable expression. The presence of numerous clones in flasks E1 - E4 together with the absence of surviving cells in control flasks C1 - C5 provides phenotypic evidence that DNA transfer takes place only by means of I-[A-CDI] and not as a result of any effects of I-A or unconjugated protein constituents. Frequencies of stable transfection resulting from ptkNEO transfer (mean = 2.0×10^{-5}) are lower than for pAL-8 (mean = 5.5×10^{-5}).

Genotypic evidence for *neo* transfer is provided in Figure 9., which shows the results of Southern hybridisation of genomic DNA samples extracted from the HepG2 cell line and from transfected populations derived from it. Figure 9A. illustrates the absence of *neo* in two mammalian cell lines including HepG2. Figure 9B. shows presence of the *neo* sequence in the transfected HepG2 populations E1 - E4, grown on G418 to high cell number through several passages, over a total period of 44 days, from duplicate flasks of those shown in Figure 8.

DISCUSSION

The carbodiimide modification of proteins to produce *N*-acylurea derivatives with DNA-binding properties and their potential in gene transfer has been reported previously [23, 39]. Here, we demonstrate that cross-linking insulin to carbodiimide-modified albumin produces a soluble macromolecule capable of both specific recognition of the insulin receptor and DNA transport, thereby affording a possible means of targeted transfection from solution without the involvement of particulate intermediates, excessive foreign chemicals, cell disruption or infective agents. We provide evidence which suggests that, by means of this DNA carrier system, the exogenous gene *neo* from the bacterial transposon Tn5 has been inserted through the insulin receptor endocytotic pathway into the HepG2 hepatoma cell line and stably expressed.

In keeping with our proposed model (Fig. 1.), the gel filtration fraction of insulin-[*N*-acylurea albumin] used for transfection was selected in order to achieve an optimal combination of protein concentration, DNA binding capacity and insulin content (Figs. 2 and 3.). Measuring DNA binding reactions by means of nitrocellulose filter and band shift assays, and using an equivalent fraction of insulin-albumin conjugate as control, we have shown that binding takes place exclusively as a result of interaction of DNA with *N*-acylurea moieties belonging to the modified albumin component of the conjugate (Figs. 3,4 and 5). In addition, we have shown that insulin-albumin conjugates bind specifically to insulin receptor sites on HepG2 cells, recognition being due to the insulin component of the conjugate, and that the presence of *N*-acylurea substituents does not adversely affect conjugate-receptor binding characteristics (Fig. 7B.). Furthermore, we have demonstrated that conjugate-DNA binding does not interfere with conjugate-receptor binding (Fig. 7C.).

The complex which arises when DNA is mixed with insulin-[*N*-acylurea albumin] is soluble under certain titration conditions as suggested by (i) reversibility of binding in the presence of high monovalent ion concentration (result not shown), and (ii) electrophoretic mobility of the complex (Figure 5, wells 5 and 6). However, as the ratio of conjugate to DNA in the mixture is raised, large electrophoretically inert complexes which may be insoluble are formed (Figure 5, wells 7 and 8). In our transfection protocol, therefore, we were careful to avoid

the possibility of insoluble material being precipitated in the medium. We achieved this by establishing conjugate-receptor binding during a pre-incubation period, following this with removal of unbound conjugate, and only then adding expression vector DNA for attachment to the receptor-bound carrier. Receptor cycling was inhibited by temperature control until the DNA binding step was complete. This approach to binding allows the addition of excess amounts of conjugate and DNA at each step, maximises binding and possibly increases transfection efficiency. The alternative would be to ensure removal of particulate material by membrane filtration of the complex before addition to cells. This might be a suitable approach for *in vivo* gene transfer studies, but further investigation into the interactions of soluble I-[A-DI]-DNA complex and components of the medium would be a desirable prerequisite.

A clear demonstration of stable expression of the *neo* gene following ptkNEO and pAL-8 transfer to HepG2 cells is provided by (i) the appearance of distinct clones among I-[A-CDI]-vector treated cells grown for 28 days on a G418 regime lethal to untreated cells (Figure 8.), and (ii) identification of the exogenous *neo* sequence in DNA extracted after clone populations produced in parallel to those shown in Figure 8. had been grown for 44 days, through a number of subculture steps, on the antibiotic (Figure 9.). Complete lack of survivors in control flasks in which cells were provided with vector in the presence of no protein, I + A, I + A-CDI and I-A respectively (Figure 8.) constitutes evidence that these gene transfer events depended on an agent capable of binding to both DNA and insulin receptor. Furthermore, free A-CDI, present at a concentration equivalent to that in the I-[A-CDI] conjugate, does not bring about transformation, indicating that non-specific cation-induced uptake is not involved. On these premises we suggest that gene transfer has taken place under the direction of insulin via the endocytotic pathway.

The size of the G418-resistant clones developed (Figure 8.) varied greatly within each treated population regardless of vector. The phenomenon has measurable parallels in plant genetics: individual *neo*-positive plantlets grown up from Ti plasmid-transformed tissue show wide quantitative difference in resistance to kanamycin; this is generally thought to indicate that expression of inserted genes is enhanced or suppressed according to genomic position, which varies as a result of random integration [40]. In enumerating clones we ignored those which were particularly small, thereby estimating transformation frequencies conservatively. Our results show I-[A-CDI]-induced stable gene transfer to be approximately 20 times (ptkNEO) and 55 times (pAL-8) greater than typical stable transformation of HepG2 cells brought about by the calcium phosphate procedure [41].

Ligand-directed passage of DNA across the plasma membrane is likely to be efficient. During subsequent intracellular processing, however, the route of DNA is less controlled, and numerous hazards in the form of enzymatic or membrane barriers

are encountered. Ultimate delivery to the nucleus must still be a rare event. The use of insulin as ligand might offer an advantage in this respect. In contrast with some ligand proteins which are delivered via membrane-bound vesicles to lysosomes, it is increasingly evident that insulin degradation inside the cell takes place in endosomal rather than lysosomal compartments [42]; thus DNA still associated with the I-[A-CDI] complex after entry may not be directed actively towards lysosomal nucleases.

Exploitation of the efficiency and specificity of cell entry afforded by receptor-mediated endocytosis has been suggested in various ways previously. For example, Poznansky *et al* have achieved delivery of α-1,4-glucosidase to muscle cells and hepatocytes using a covalently conjugated enzyme-albumin-insulin complex [31], and Cheng *et al* have devised a mechanism of covalent attachment between nucleic acid and protein with the purpose in view of genetic correction via endocytosis [43]. The method of gene transfer presented here combines the elegance of ligand-directed carrier insertion with non-covalent reversible binding of the DNA to the carrier [23], a possible advantage in the cell interior. Recent work by Wu and Wu published during the course of our investigations illustrates use of a parallel approach: by employing as carrier a conjugate of asialo-orosomucoid and high molecular weight cationic poly-L-lysine, which binds DNA electrostatically in much the same manner as *N*-acylurea protein, those workers have demonstrated transient expression of the bacterial *cat* gene in HepG2 hepatoma cells in culture [44, 45] and in rat liver *in vivo* [46]. The work we describe in the present communication adds to the growing body of evidence that receptor mediated endocytosis provides a pathway by which exogenous molecules, including DNA, may be directed into the interior of target cells, and in particular shows that stable gene transformation may result from the use of this facility.

ACKNOWLEDGEMENTS

The authors thank Professor Peter Jones (Comprehensive Cancer Center, University of Southern California School of Medicine) for his encouragement in this work and for his courtesy in providing pAL-8, Dr. Shirley Taylor (Division of Biology, California Institute of Technology) for ptkNEO (constructed by Dr. Barbara Wold of the same institution), Professor Dave R. Woods (Microbial Genetics Research Unit, University of Cape Town) for *Escherichia coli* HB101, Professor Gerry Coetzee (Department of Medical Biochemistry, University of Cape Town) and Dr. Denver Hendricks (Department of Biochemistry, University of the Western Cape) for HepG2 cells and related technical help, and Mrs. Lorraine van Hooff (University of Durban-Westville) for expert photographic assistance.

This work was jointly supported by the Foundation for Research Development and University of Durban-Westville Research Funds.

REFERENCES

1. M. J. Gething and J. Sambrook, *Nature (London)* **293**, 620 (1981).

2. H. R. Scholer and P. Gruss, *Cell* **36**, 403 (1984).

3. J. E. Dick, M. C. Magli, D. Huszar, R. A. Phillips and A. Bernstein, *Cell* **42**, 71 (1985).

4. B. Hjelle, E. Liu and J. M. Bishop, *Proc. natn. Acad. Sci. USA* **85**, 4355 (1988).

5. R. Jaenisch, *Science* **240**, 1468 (1988).

6. F. Costantini, K. Chada and J. Magram, *Science* **233**, 1192 (1986).

7. F. L. Graham, S. Bacchetti and R. McKinnon, in *Introduction of macromolecules into viable mammalian cells* (Eds. R. Baserga, C. Croce and G. Rovera), p. 3. Alan R. Liss, Inc., New York (1980).

8. S. Kawai and M. Nishizawa, *Mol. cell Biol.* **4**, 1172 (1984).

9. R. S. Kerbel, C. Waghorne, M. S. Man, B. Elliott and M. L. Breitman, *Proc. natn. Acad. Sci. USA* **84**, 1263 (1987).

10. R. Pine, D. E. Levy, N. Reich and J. E. Darnell, Jr., *Nucleic Acids Res.* **16**, 1371 (1988).

11. M. Graessmann and A. Graessmann, *Methods in Enzymology* **101**, 482 (1983).

12. H. Stopper, H. Jones and U. Zimmermann, *Biochim. biophys. Acta* **900**, 38 (1987).

13. M. Fechheimer, J. F. Boylan, S. Parker, J. E. Sisken, G. L. Patel and S. J. Zimmer, *Proc. natn. Acad. Sci. USA* **84**, 8463 (1987).

14. C-J. Tsao, T. Hosoi, H. Hirai, T. Saito, T. Okabe, A, Urabe and F. Takaku, *Int. J. cell Cloning* **5**, 220 (1987).

15. F. C. Wiberg, P. Sunnerhagen, K. Kaltoft, J. Zeuthen and G. Bjursell, *Nucleic Acids Res.* **11**, 7287 (1983).

16. A. Vainstein, A. Razin, A. Graessmann and A. Loyter, *Methods in Enzymology* **101**, 492 (1983).

17. A. Cudd and C. Nicolau, in *Liposome Technology, vol. II.* (Ed. G. Gregoriadis), p.207. CRC Press, Cleveland, Ohio (1984).

18. P. A. M. Peeters, C. A. M. Claessens, W. M. C. Eling and D. J. A. Crommelin, *Biochem. Pharmacol.* **37**, 2215 (1988).

19. A. D. Miller, M-F. Law and I. M. Verma, *Mol. cell Biol.* **5**, 431 (1985).

20. H. Varmus, *Science* **240**, 1427 (1988).

21. J. L. Goldstein, R. G. W. Anderson and M. S. Brown, *Nature (London)* **279**, 679 (1979).

22. R. Timkovitch, *Analyt. Biochem.* **79**, 135 (1977).

23. B. Huckett, H. Gordhan, R. Hawtrey, N. Moodley, M. Ariatti and A. Hawtrey, *Biochem. Pharmacol.* **35**, 1249 (1986).

24. B. B. Knowles, C. C. Howe and D. P. Aden, *Science* **209**, 497 (1980).

25. E. H. Harley, C. L. Oliver, L. Rhodes-Harrison, R.T. Mew, G. Lecatsas and W. du T. Naude, *S. A. J. Science* **78**, 112 (1982).

26. T. Maniatis, E. F. Fritsch and J. Sambrook, *Molecular cloning, a laboratory manual*, Cold Spring Harbour Laboratory Cold Spring Harbour, NY (1982).

27. L. Katz, D. Kingsbury and D. Helinski, *J. Bact.* **114**, 577 (1973).

28. R. Radloff, W. Bauer and J. Vinograd, *Proc. natn. Acad. Sci. U.S.A.* **57**, 1514 (1967).

29. P. W. J. Rigby, M. Dieckmann, C. Rhodes and P. Berg, *J. molec. Biol.* **113**, 237 (1977).

30. A. P. Feinberg and B. Vogelstein, *Analyt. Biochem.* **137**, 266 (1984).

31. M. J. Poznansky, R. Singh, B. Singh and G. Fantus, *Science* **223**, 1304 (1984).

32. O. H. Lowry, N. J. Rosebrough, A. L. Farr and R. J. Randall, *J. biol. Chem.* **193**, 265 (1951).

33. U. K. Laemmli, *Nature (London)* **227**, 680 (1970).

34. C. Shih and R. A. Weinberg, *Cell* **29**, 161 (1982).

35. E. M. Southern, *J. molec. Biol.* **98**, 503 (1975).

36. S. Gammeltoft, *Physiol. Revs.* **64**, 1321 (1984).

37. S. Marshall, *J. biol. Chem.* **260**, 4136 (1985).

38. F. Colbere-Garapin, F. Horodniceanu, P. Kourilsky and A-C. Garapin, *J. mol. Biol.* **150**, 1 (1981).

39. M. Ariatti and A. O. Hawtrey, *Med. Hypoth.* **24**, 29 (1987).

40. S. G. Rogers, R. B. Horsch and R. T. Fraley, *Methods in Enzymology* **118**, 627 (1986).

41. G. J. Darlington, *Methods in Enzymology* **151**, 19 (1987).

42. O. Sonne, *Physiol. Revs.* **68**, 1129 (1988).

43. S-Y. Cheng, G. T. Merlino and I. H. Pastan, *Nucleic Acids Res.* **11**, 659 (1983).

44. G. Y. Wu and C. H. Wu, *J. biol. Chem.* **262**, 4429 (1987).

45. G. Y. Wu and C. H. Wu, *Biochemistry* **27**, 887 (1988).

46. G. Y. Wu and C. H. Wu, *J. biol. Chem.* **263**, 14621 (1988).

LEGENDS

Figure 1. (Figure 1.1 in thesis)

Binding interactions of a conjugate consisting of a small polypeptide ligand such as insulin covalently cross-linked to N-acylurea albumin. The ligand component retains its facility for specific receptor binding. The *N*-acylurea albumin component, produced by carbodiimide modification of serum albumin, bears numerous positively charged *N*-acylurea moieties (⊕) in place of acidic amino acid side chain carboxyl groups (aspartate and glutamate residues) and readily interacts electrostatically with negatively charged DNA phophodiester backbone (⊖). I, insulin; A-CDI, *N*-acylurea albumin; α and ß, insulin receptor subunits.

Figure 2. (Figure 2.6 in thesis)

Gel filtration chromatography of insulin, albumin and derivatives. Proteins were purified and characterised by gel filtration through Sephadex G-100 in a 7 X 1500mm column using 10mM acetate buffer, pH 4.0, containing 5%v/v DMSO as eluant at a flow rate of 6.0ml/h. Absorption at 280nm was determined as a continuous trace, and fractions of 1ml collected. *A.* Unmodified insulin (I), unmodified albumin (A) and *N*-acylurea albumin (A-CDI). *B.* Insulin-albumin (I-A) and insulin-[*N*-acylurea albumin] (I-[A-CDI]). Reaction mixtures were dialysed exhaustively prior to loading on to the column; reaction mixtures containing insulin were dialysed in low molecular weight cut-off dialysis tubing. The horizontal bar shows those conjugate fractions containing the highest proportions of insulin to total protein as estimated by immuno-dot-blot analysis.

Figure 3. (Figure 3.3 in thesis)

DNA binding capacity of insulin-albumin conjugates fractionated by gel filtration. DNA binding was assayed by the nitrocellulose filtration technique. Each assay mixture contained 4.5ng [^{3}H]pBR322, 5.5ng unlabelled pBR322 and 1.5µg protein conjugate. Absorption at 280nm of I-A and I-[A-CDI] (——); DNA bound by I-A (heavy hatching); DNA bound by I-[A-CDI] (light hatching). The horizontal bar indicates those fractions pooled for use in subsequent binding and transfection experiments.

Figure 4. (Figure 3.4B in thesis)

Binding of DNA to insulin-[N-acylurea albumin] at varying protein concentration. DNA binding was assayed by the nitrocellulose filtration technique. Each assay mixture contained 4.5ng [^{3}H]pBR322 and 5.5ng unlabelled pBR322. Each protein conjugate was added in the form of a pooled gel filtration peak component as indicated in Figure 3. I-[A-CDI] (●); I-A control (○).

Figure 5. (Figure 3.2B in thesis)

Band shift assay of binding interactions between insulin-albumin conjugates and DNA. pBR322 was incubated with protein conjugate at 20°C for 30min prior to agarose gel electrophoresis. Each incubation mixture contained 50mM NaCl, 20mM Tris-HCl (pH 7.0), 1mM EDTA, 0.45µg pBR322, and a variable amount of protein conjugate as indicated, in a final volume of 10µl. 1, pBR322 control (no protein); 2 and 3, pBR322 plus I-A (2.0µg and 3.3µg respectively); 4-8, pBR322 plus I-[A-CDI] (0.4µg, 1.0µg, 1.66µg, 2.66µg, and 3.3µg respectively).

Figure 6. (Figure 4.1 in thesis)

Binding of human insulin to HepG2 insulin receptors. Semi-confluent monolayers of HepG2 cells were grown and prepared for binding as described in Materials and Methods. A. [^{125}I-Tyr^{A14}] human insulin (2µl, 0.4ng, 2.36 X 10^{5} cpm) was added to each well and the incubation period at 10°C varied. B. Variable amounts of [^{125}I-Tyr^{A14}] human insulin were added to the wells and incubation at 10°C allowed to continue for 18 hours. After incubation each well was processed as described in Materials and Methods.

Figure 7. (Figure 4.2 in thesis)

Competitive displacement of human insulin from HepG2 insulin receptors by porcine insulin and its conjugated derivatives. Semi-confluent monolayers of HepG2 cells were grown and prepared for binding as described in Materials and Methods. [^{125}I-Tyr^{A14}] human insulin (2µl, 0.4ng, 2.36 X 10^{5} cpm) was added to each well together with variable amounts of unlabelled competitor protein as shown. A. Porcine insulin. B. (Porcine) insulin-albumin conjugates I-A (Φ) and I-[A-CDI] (●), plus albumin controls (○). C. I-[A-CDI] which had been incubated with pBR322 DNA in binding buffer at 20°C for 30 minutes prior to addition (●), plus DNA controls (○); DNA values underlined. Incubation was at 10°C for 18 hours. Subsequent processing was as described in Materials and Methods.

Figure 8. (Figure 5.7 in thesis)

Appearance of clones resistant to G418 following insulin-[N-acylurea albumin] mediated transfer of the expression vectors ptkNEO and pAL-8 to HepG2 cells. Cells were subjected to transfection and G418 selection as described in Materials and Methods. After 28 days of G418 treatment, one complete experimental set of flasks was stained for the visualisation of clones as follows: medium was removed, cells rinsed twice with PBS (5ml), fixed with redistilled methanol (5ml, 2min), stained with a 1/10 aqueous dilution of standard Giemsa (5ml, 2-5min), then rinsed generously with water several times. Flasks were drained and air-dried.
TRANSFECTION FLASKS: E1 and E2, I-[A-CDI] (40µg) plus ptkNEO (4µg); E3 and E4, I-[A-CDI] (40µg) plus pAL-8 (4µg).
TRANSFECTION CONTROL FLASKS: C1, no protein, no vector; C2, no protein, pAL-8 (4µg); C3, unconjugated I (18.25µg) and A (21.75µg) plus pAL-8 (4µg); C4, unconjugated I (18.25µg) and A-CDI (21.75µg) plus pAL-8 (4µg); C5, I-A (40µg) plus pAL-8 (4µg).

Figure 9. (Figure 5.8 in thesis)

Genomic Southern hybridisation: detection of the neo sequence in stably G418 resistant HepG2 cells following insulin-[N-acylurea albumin] mediated transfection. Cells were subjected to transfection and G418 selection as described in Materials and Methods. After 28 days of G418 treatment, one complete experimental set of flasks was stained for the visualisation of clones (Figure 8.). Flasks of the duplicate set containing surviving cells (E1 - E4) were trypsinised and each of the cell populations propagated through several passages on a G418 maintenance regime for a further 16 days until there were sufficient cells for genomic DNA extraction. DNA was analysed for the *neo* sequence using as probe a 923bp PstI fragment of ptkNEO within the *neo* gene.
A. MINI GEL BLOT : 1, PstI digested ptkNEO (2ng) plus salmon sperm carrier DNA (3µg); 3, *neo* fragment excised from PstI digested ptkNEO (923bp) (2ng) plus carrier DNA (3µg); 5, PstI digested HepG2 cell line DNA (12µg); 6, PstI digested HeLa cell line DNA (12µg).
B. LARGE GEL BLOT: 1, undigested ptkNEO (4913bp) (1ng) plus carrier DNA (7µg); 3, undigested pAL-8 (>10000bp) (1ng) plus carrier DNA (7µg); 5, *neo* fragment excised from PstI digested ptkNEO (923bp) (1ng) plus carrier DNA (7µg); 7-10, PstI digested DNAs from G418 resistant HepG2 cell populations E1 - E4 respectively (28µg per well).

THE BRIGHT FIELD

I have seen the sun break through
to illuminate a small field
for a while, and gone my way
and forgotten it. But that was the pearl
of great price, the one field that had
the treasure in it. I realise now
that I must give all that I have
to possess it. Life is not hurrying

on to a receding future, nor hankering after
an imagined past. It is the turning
aside like Moses to the miracle
of the lit bush, to a brightness
that seemed as transitory as your youth
once, but is the eternity that awaits you.

R.S. Thomas

From *Laboratories of the Spirit*
In *Later Poems 1972-1982*
Macmillan, London, 1983.