Modified Virus

The present invention relates to novel recombinant viruses suitable for use in gene therapy. Such recombinant viruses exhibit an altered tropism conferred by incorporation of one or more non-native polypeptides into one or more viral components, or the replacement of such components by non-native polypeptides. These non-native polypeptides can comprise elements which mimic the structure of the original viral component so as to permit inclusion of the polypeptide in functional virion particles. These non-native polypeptides can also comprise elements which confer a non-native ligand binding function (i.e. altered tropism) to such virion particles. Essentially, the non-native polypeptides employed in the invention have primary and secondary amino acid structures that enable their correct folding in the nucleus or cytosol of mammalian host cells and their subsequent transport through the nuclear membrane. Such structures permit the assembly of functional recombinant virion particles with an intact non-native binding function and thus an altered tropism.

Clinical gene therapy was introduced in 1989. The aim at that time was to correct gene defects in the immune system via the *in vitro* introduction of a healthy gene into the defective cells of the patient and transfusion of the treated cells back to the patient. Since that time, the indications and possible molecular uses of gene therapy have increased dramatically. Today, ten years after its introduction, one can envisage the use of gene therapy to treat e.g. vascular diseases, cancer, inflammatory diseases and infectious diseases such as HIV.

At present, however, gene therapy is still not a useful

method in human medicine. One main reason is that gene therapy demands the packaging of the genes to be delivered into gene-carriers, or vectors, which can be administered to patients and which will target the genes only to the intended cells. Such vectors have so far not been available in reliable form.

Several classes of viruses have been considered as vectors for gene therapy applications, the most commonly used being adenoviruses, retroviruses, lentiviruses and adeno-associated viruses.

Ideal vectors for gene therapy would be those which can be administered systemically and yet deliver the desired genetic material specifically to desired cells or tissues. However, the viruses currently considered for human gene therapy applications have a broad tropism, being able to infect many different types of cells in the human body. This limits the potential safety of viral vectors and prohibits their use for systemic administration. For this reason, the development of targeted virus vectors, capable of infecting only selected cells, has been described as the holy grail of gene therapy.

One of the most widely investigated group of vectors for gene therapy applications is the adenoviruses. Human adenoviruses are divided into six hemagglutination groups (A-F) and each hemagglutination group is further subdivided into different serotypes. All in all there exist more than 40 different human adenovirus serotypes. The adenovirus which has been most frequently used for human gene therapy is adenovirus type 5 (Ad5) which belongs to hemagglutination group C.

Adenoviruses (Ad) are DNA viruses without an envelope, shaped as regular icosahedrons with a diameter of 60-85

nm.

The adenoviral capsid comprises 252 capsomeres, 240 hexons and 12 pentons (Ginsberg et al., Virology, 28, 782-83 (1966)). The hexons and pentons are derived from three viral proteins. The hexon comprises three identical proteins of 967 amino acids each, namely polypeptide II. The penton contains a base, which is bound to the capsid, and a fiber, which is non-covalently bound to, and projects from, the penton base. Proteins IX, VI and IIIa also are present in the adenoviral coat and are thought to stabilise the viral capsid.

Cell binding takes place through the fiber proteins, anchored to the virion at the vertices of the icosahedron. The fiber protein is not necessary for assembly and release of intact virions. Assembly of virions take place in the nucleus of infected cells.

The fiber protein, which is a homotrimer of a fiber polypeptide (namely adenoviral polypeptide IV), contains three functionally different parts: an N-terminal tail anchoring the fiber non-covalently to the penton base in the virion and which furthermore contains the nuclear-localization signal; a shaft domain comprising a variable number of repeats of a ~15 amino acid fiber shaft motif, (e.g. which is repeated six times in Ad3 and 22 times in Ad2 and Ad5); and a C-terminal globular domain, the knob, which contains the ligand which binds to the cellular Ad-receptor (Chrobozek et al., Microbiology and Immunology, (1995), p. 163-200). The knob is also functionally responsible for fiber trimerisation (i.e. it incorporates a trimerisation motif).

Each shaft repeat has two three-amino acid regions which

form beta-sheets and two amino acid regions which constitute the turns of the native extended fiber shaft. The crystal structure of the trimerised, cell binding domain has been determined and shows a unique topology different from other anti-parallel β-sandwiches (Xia et al., Structure 2: 1259-1270, (1993)). Binding of the fiber to the penton base of the virion can take place also in a cell-free system, i.e. the fiber can bind to fiberless virions (Boudin et al., Virology, 116: 589-604, (1982)).

Efforts to modify adenovirus fibers (e.g. produce recombinant fiber proteins) in order to modify the properties of adenovirus vectors, for example by altering tropism or cell binding, have been made and have been reported in the prior art.

The adenovirus fiber protein performs several biological functions which must be retained in order to produce active virus particles. The following fiber features are deemed to be of key importance in the construction of functional modified fiber proteins:

i) the ability to form parallel homotrimers. This function is mediated by the N-terminal amino acid sequence of the wild type fiber knob.

ii) the ability to bind to the penton base to form penton capsomeres. This function is mediated by the wild type fiber tail.

iii) the ability to express a cell binding ligand allowing for attachment to target cells. In the native situation, this function is mediated by the wild type fiber knob (which binds to the cellular Ad-receptor).

iv) the capability of transport into the nucleus of infected cells, which is vital to virus formation. This function is mainly, but perhaps not exclusively, mediated by the wild type fiber tail.

Previous attempts to change the tropism of adenovirus have involved genetic modification of fibers and knobs. However, this approach has not proved to be very successful. A major problem has been the incorporation of novel ligands in a functional context, which are capable of changing the tropism without interfering with the trimerisation of the fiber. For example, a short peptide ligand has been added C-terminally of the knob (Michael et al., Gene Therapy, 2: 660-8, (1995)) and a nonapeptide has been introduced into one of the knob "loops" (Dmitriev et al., J. Virol., 72(12):9706-9713 (1998)). However, the knob has a very complex structure due to interactions between the three fiber subunits which are necessary to conserve cell binding and trimerisation. Therefore, the knob only tolerates insertion of a few amino acids and no general method for construction of functional, genetically re-targeted adenovirus fibers exists to date.

Attempts have also been made to introduce new ligands into other parts of the adenovirion. By introducing the FLAG tetra-amino acid motif into the Ad penton, it has been shown possible to target Ad to cells normally not infected by Ad. The re-targeting was achieved by targeting with a bi-specific antibody where one specificity was directed against the FLAG peptide and the other against the new target cell (Wickham et al., J. Virol., 70: 6831-6838, (1996)).

A previously unaddressed problem encountered in the production of efficient recombinant viruses for gene therapy, is that of ensuring the functional folding of recombinant components upon expression in the nucleus and cytosol of host cells. This is particularly relevant where the wild-type virus to be engineered employs component expression in these intracellular locations during replication, for example, as in

adenovirus. Folded protein structures, particularly those which rely on disulphide bonds in cysteine bridges to maintain a functionally correct conformation, for example those derived from antibodies, can be rendered mis-folded and non-functional when expressed in the reducing environment of the nucleus and cytosol as part of a recombinant viral component. Thus, there exists in the art an unsatisfied and previously unappreciated need for protein structures which can retain a functional conformation when expressed as viral components in these cellular compartments, such that the binding functions which the structures comprise or support retain a binding function, particularly in the absence of the ligand for which binding is desired.

There are therefore major problems associated with the genetic engineering of Adenovirus fibers useful in the construction of recombinant re-targeted adenovirus (Ad-virus) for human gene therapy. These problems are deemed to be particularly important since more patients have been treated with adenovirus vectors than with any other type of vector (Trapnell etal., Biotechnology, 5: 617-625, (1994)). The present invention is directed towards circumventing such problems in the construction of genetically re-targeted viruses for gene therapy, where the new viral tropism has been accomplished by the introduction of a new cell binding ligand into a viral component protein.

SUMMARY OF THE INVENTION

In a first aspect of the present invention there is provided a modified virus comprising one or more non-native polypeptides, which polypeptide comprises one or more framework moieties which possesses a folded three dimensional structure, each framework moeity containing one or more binding moieties, which polypeptide is capable of being expressed in the cytoplasm and nucleus

of a mammalian host cell in a conformation which is maintained in the absence of a ligand for said binding moieties, said conformation allowing said binding moieties subsequently to bind with said ligand, and which polypeptide is capable of transport though the nuclear membrane, wherein said modified virus has an altered tropism conferred by said binding moieties.

In a second aspect of the present invention there is provided a modified viral protein comprising a non-native polypeptide as defined above.

In a third aspect of the present invention there is provided a cell containing a modified virus or viral protein as defined above.

In a fourth aspect of the present invention there is provided a permissive cell for a modified virus, the modified virus as defined above, which cell is capable of being cultured to propagate said modified virus.

In a fifth aspect of the present invention there is provided a method for producing a modified virus, the modified virus as defined above in cell culture, comprising the steps of:

i) genetically modifying a virus to produce a modified virus comprising one or more non-native polypeptides, which polypeptide comprises one or more framework moieties each containing one or more binding moieties, which polypeptide is capable of being expressed in the cytoplasm and nucleus of a mammalian host cell in a conformation which is maintained in the absence of a ligand for said binding moieties, said conformation allowing said binding moieties subsequently to bind with said ligand, and which polypeptide is capable of transport though the nuclear membrane, wherein said modified virus

has an altered tropism conferred by said binding moieties;

ii) infecting permissive cells with said modified virus;

iii) culturing said cells to produce the virus; and

iv) harvesting, and optionally, purifying the modified virus produced.

In a sixth aspect of the present invention there is provided a method of regulating the replication of a modified virus, the modified virus as defined above comprising the steps of:

i) constructing a modified virus such that a cleavage site is positioned between a binding moiety required for cell infection and the remainder of the recombinant viral component of which the binding moiety forms part; and,

ii) bringing said recombinant virus into contact with a cleavage agent or cleavage means capable of cleaving said binding moiety from said viral component and thereby preventing the recombinant virus from undergoing further infection cycles.

In a seventh aspect of the present invention there is provided a method of determining the suitability of a non-native polypeptide for use in the preparation of a viral vector by determining its solubility in a cell system.

In an eighth aspect of the present invention there is provided a method of assaying the solubility of a non-native polypeptide as defined above or a modified viral protein as defined in the second aspect of the present invention, which method comprises the steps of:

i) expressing said non-native polypeptide or a viral component protein comprising said non-native polypeptide in permissive cells;

ii) subjecting the cells to lysis to produce a cell lysate.

iii) separating the soluble and insoluble fractions of the cell lysate.

iv) analysing the soluble and insoluble fractions of the cell lysate for the content of said non-native polypeptide or viral component protein comprising said non-native polypeptide; and

v) comparing the relative content of said non-native polypeptide or viral component protein comprising said non-native polypeptide in the soluble and insoluble fractions.

In a ninth aspect of the present invention there is provided a use of the modified virus as defined above in the preparation of a medicament for the treatment of tumour cells or proliferating cells.

In a tenth aspect of the present invention there is provided a pharmaceutical composition comprising a modified virus as defined above and a pharmaceutically acceptable carrier or excipient.

DETAILED DESCRIPTION OF THE INVENTION

The invention described herein relates in part to the production of functional recombinant adenoviral fiber proteins with a new tropism facilitated by removing or ablating (e.g. blocking or inactivating) the native cell binding domain and either replacing it with, or adding, a non-native polypeptide comprising an external cell binding ligand and an external trimerisation motif.

Curiel et al. in WO 99/41359, Wickham et al. in WO 98/54346 and Spooner and Epenetos in US 5,885,808 disclose various modifications to adenoviral components with the aim of producing recombinant viruses with an altered tropism. None of the these disclosures provide a solution to, or indeed address or even recognise, the problems associated with the expression of functional non-native viral components in the nucleus and cytosol of host cells, a problem solved by the present invention.

Thus, in one aspect the invention provides a modified virus comprising a non-native polypeptide, which polypeptide comprises one or more framework moieties each containing one or more binding moieties, which polypeptide is capable of being expressed in the cytoplasm and nucleus of a mammalian host cell in a conformation which is maintained in the absence of a ligand for said binding moieties, said conformation allowing said binding moieties subsequently to bind with said ligand, and which polypeptide is capable of transport though the nuclear membrane, wherein said modified virus has an altered tropism conferred by said binding moieties.

A "modified" virus according to the invention is thus a virus which differs to a native (i.e. wild-type) virus. Generally speaking, the virus is modified such that a component of the virus is altered structurally over a native or wild-type (i.e. unmodified) component, or such that a structural component or feature is added to the virus, which is not present in the native or wild-type form. Advantageously, according to the present invention, a property or behaviour of the component is altered. In other words, the modified virus, differs functionally over an unmodified (native/wild-type) virus (e.g. by exhibiting an altered tropism). As discussed

above (and also further below), the virus is conveniently modified using genetic engineering techniques. Accordingly, a modified virus according to the present invention is advantageously a recombinant virus.

A modified virus according to the invention can be derived from any virus, and in particular any virus which may be used as the basis of a viral vector for gene therapy. Representative viral families include adenoviruses, retroviruses, lentiviruses and adeno-associated viruses and particularly include members of the family Adenoviridae or other virus families where viral structural components are synthesized and/or assembled in the nucleus or cytosol of the host cell, such as Reoviridae, Picornaviridae, Parvoviridae, Papovaviridae and Caliciviridae. In a preferred aspect, modified viruses of the invention are modified forms of adenoviruses, in particular Human adenoviruses and more particularly Human adenovirus type 5.

"Non-native polypeptide" as defined herein is a polypeptide sequence of two or more amino acids, the complete sequence of which is not found in a functionally equivalent position in the amino acid sequence of the wild-type virus, or more particularly in the wild-type viral component protein to be engineered.

In certain preferred embodiments of the invention, the "non-native polypeptide" may also be non-native in the sense of not occurring in nature, i.e. being a synthetically or artificially constructed or prepared polypeptide.

According to the present invention, the conformation of the non-native polypeptide after expression in the intracellular environment and, more importantly, also

its conformation in the extracellular environment, should be such that it is capable of binding to a desired ligand, e.g. a receptor or other molecule expressed on the outer surface of a target cell for the modified virus.

As mentioned above, advantageously according to the present invention, the non-native polypeptide, when expressed in the cytoplasm of a mammalian cell has a conformation which may be maintained in the absence of a ligand for the binding moiety(ies) of the polypeptide. Such a conformation is thus a stable conformation. In other words, upon expression in the cytoplasm, the polypeptide assumes a conformation which is maintained in the cytoplasm, and when the polypeptide is transported into the nucleus and incorporated or assembled into a viral particle. Moreover, the conformation is such that the binding moiety or moieties of the polypeptide is or are able to bind to their ligand, when exposed to it.

Thus, an important feature of the non-native polypeptide of the invention is that it is able to fold correctly in the cytoplasm or cytosol of a mammalian cell i.e. assume a tertiary or three-dimensional structure which permits the functionality of the binding moiety(ies) to be retained (i.e. that the binding moieties retain their binding activity towards their ligand). In other words, the binding moiety remains capable of functional binding to its ligand.

The solubility of recombinant proteins is considered a good indicator of their correct folding. Therefore non-native polypeptides of the invention may advantageously be chosen not only on the functionality of binding moieties, stability of conformation and ability to form part of an assembled virion, but also on the basis of

solubility within cells in which a non-native polypeptide, or a viral component protein containing or comprising a non-native polypeptide, are expressed. For example, recombinant fiber proteins may be tested for their total expression in eukaryotic (e.g. insect) cells and the proportion recovered in the soluble fraction of the cell lysates determined.

The Inventors have found that phenotypic analysis of fiber-ligand fusion constructs expressed as recombinant proteins in baculovirus-infected insect cells showed that their degree of solubility correlated with the recovery of viable recombinant adenovirus. Therefore, it is preferred that any non-native polypeptide of the invention, or a viral component protein containing or comprising a non-native polypeptide, for example a fiber-ligand fusion construct, is characterised with regard to its solubility and target cell attachment as recombinant protein, prior to the re-insertion of the corresponding gene into the viral, for example adenovirus genome.

Thus the invention also provides a modified virus wherein said non-native polypeptide is selected using a solubility assay of the non-native polypeptide or the viral component protein comprising the non-native polypeptide. The non-native polypeptide is selected such that greater than 25%, preferably greater than 30%, more preferably greater than 50% or even 70%, of the non-native polypeptide or the viral component protein comprising the non-native polypeptide is present in the soluble fraction of cell lysates of cells expressing the non-native polypeptide or viral component protein comprising the non-native polypeptide. A suitable assay is described in the examples herein which may be applied *mutatis mutandis* to other cell expression systems. Reference herein to a non-native polypeptide or a viral

component comprising said non-native polypeptide being soluble in the cellular environment should be interpreted in line with these limits, i.e. when at least 25-30% of the polypeptide is present in the soluble fraction of a cell lysate, the polypeptide can be considered 'soluble'.

Thus, according to a further aspect, the present invention provides a method of determining the suitability of a non-native polypeptide (e.g. a recombinant viral fusion protein) for use in the preparation of a viral vector by determining its solubility in a cell system. More specifically is provided a method of assaying the solubility of a non-native polypeptide or a modified viral protein component of the invention, comprising the steps of i) expressing said non-native polypeptide or a viral component protein comprising said non-native polypeptide in permissive cells; ii) subjecting the cells to lysis to produce a cell lysate; iii) separating the soluble and insoluble fractions of the cell lysate; iv) analysing the soluble and insoluble fractions of the cell lysate for the content of said non-native polypeptide or viral component protein comprising said non-native polypeptide; and, v) comparing the relative content of said non-native polypeptide or viral component protein comprising said non-native polypeptide in the soluble and insoluble fractions to determine the solubility.

A ligand for a binding moiety of the invention may be regarded as the ligand which corresponds to the binding moiety, or in other words, the ligand to which the binding moiety was designed or selected to bind. The ligand is thus capable of binding to the binding moiety. The ligand and its binding moiety may thus be regarded as members of an affinity-binding pair, the ligand being

a binding partner for the binding moiety.

The binding moiety thus has a binding specificity for, or is capable of binding specifically to, a desired ligand. By "binding specificity" or "binding specifically" it is meant that the binding moiety is capable of binding to the desired (or "target") ligand in a manner which is distinguished from the binding to non-target molecules or ligands. Thus, the binding moiety either does not bind to non-target molecules or exhibits negligible or substantially-reduced (as compared to the target ligand) e.g. background, binding to non-target molecules. The binding moiety thus specifically recognises the target ligand.

This specificity of binding of the binding moiety thus permits the modified virus to be selectively targeted. In other words, the binding moiety may be designed or selected to enable the modified virus to bind to a desired or "target" cell. A binding moiety may be designed or selected which binds to a ligand expressed by or on a target cell, e.g. on the cell surface.

The ligand may thus be any desired ligand, and advantageously will be a molecule expressed on the surface of a target cell, in which it is desired to achieve expression of the modified virus. The ligand may thus be a cell surface receptor, or a cell-surface antigen. The ligand may conveniently be a protein or polypeptide molecule, but may be of any molecular nature, for example a lipid or carbohydrate.

In this manner, the modified virus may be modified to bind to a desired target cell to which the native (wild-type) virus from which it is derived does not bind, or it may be modified to have a more restricted binding specificity than the wild-type virus, in other words, to

bind to only a selected or particular sub-set or type of target cell from among a broader population of target cell types to which the wild-type virus binds. The tropism of the virus is thus altered. Hence, by "altered tropism" it is meant that the modified virus exhibits a target cell binding specificity which is altered, or different, to that of the wild-type virus from which it is derived.

Generally, a non-native polypeptide according to the invention comprises at least one framework moiety and one or more binding moieties and is capable of interacting with other viral components to form a functional and infective virion particle. The ability to form part of a functional virion particle can be facilitated by the presence of one or more sequences within the non-native polypeptide that effect binding to other viral components by the non-native polypeptide itself, or a viral component in which the non-native polypeptide is comprised, in the virion assembly process.

A non-native polypeptide of the invention may replace or be incorporated into any viral component protein which is capable of interacting with a target cell. The modified (e.g. recombinant) viral component possesses a cell binding function either by the nature of a retained native structure of the viral component itself or, by a new structure conferred upon it by the incorporation of the non-native polypeptide and any structures comprised therein. In a preferred aspect of the invention, the non-native polypeptide is introduced or incorporated into or forms a fusion protein with, a viral protein component of the wild type virus, for example, an adenoviral fiber protein, and especially preferably it is incorporated such that the wild-type fiber knob (or at least the cell-binding domain thereof) is removed

(i.e. replaced).

However, in an alternative embodiment, a wild-type fiber knob (or cell-binding domain thereof) may be retained, and a further, or additional, cell-binding domain may be added by virtue of the binding moiety(ies) of the non-native polypeptide. It is a feature of the invention that the cell binding functions are altered from those of the wild type virus to be engineered, such that an altered viral tropism is present in the modified virus. Where the wild-type fiber knob (or cell-binding domain thereof) is retained, this altered tropism may be achieved (or facilitated) by modifying the wild-type knob/cell-binding domain thereof to inactivate or block the native or wild-type cell binding function.

As will be discussed further below, it may in certain circumstances be desirable to construct or engineer a modified virus in which control (e.g. temporal control) may be exhibited on the expression of the altered tropism. Thus, in such a modified virus, a native or wild-type tropism may be retained, by retaining the wild-type fiber knob or cell-binding domain thereof, in addition to providing the non-native polypeptide which confers altered tropism. For example, this is desirable for the propagation of a modified virus in conventional cell lines known in the art via the additional presence in the modified virus of a wild type binding function, for example through the controlled expression of a wild type adenoviral fiber gene from an inducible control element (e.g. promoter).

The wild-type tropism may be ablated when desired by controlling (e.g. preventing) expression of the additional wild type viral component genes, or after expression, by removing or inactivating the wild type knob/domain, in such a manner that altered tropism conferred by the non-native polypeptide may be

expressed. An "altered tropism" according to the invention thus includes a "potential" altered tropism, i.e. the potential to express an altered tropism. It also includes an altered tropism which is additional to a wild-type tropism.

As mentioned above, the modified virus of the invention is preferably prepared using genetic engineering techniques and in preferred embodiments of the invention, the non-native polypeptide is provided as part of a fusion protein with a viral protein component, preferably as a fusion protein with an adenoviral fiber protein. Such a fusion protein, or more generally, such a modified viral component protein, represents a separate aspect of the present invention. In a preferred embodiment of this aspect of the invention, the modified viral component is a modified adenoviral fiber protein comprising a non-native polypeptide as defined above.

Techniques for preparing such non-native polypeptides and introducing them into viruses or viral components are well known in the art and widely described in the literature. Thus, for example, molecular biology or genetic engineering techniques are readily available, to prepare or construct genetic sequences capable of being expressed as a modified virus, or viral component, according to the present invention. As described further in the Examples below, a nucleic acid molecule or nucleotide sequence encoding a viral component protein, such as the adenoviral fiber protein, may be modified so as to introduce a nucleotide sequence encoding the non-native polypeptide, for example so as to encode a fusion protein comprising all or part of an adenoviral fiber protein and the non-native polypeptide.

Depending upon the viral component protein into which

the non-native polypeptide of the invention is incorporated, or which component(s) it replaces, a non-native polypeptide according to the invention can optionally comprise a further element which mimics the native structure or function of the viral component (i.e. is a functional equivalent of the viral component) so as to facilitate the assembly of functional virion particles. In the case of modified adenoviral fibers according to the invention, the functional integrity of the adenoviral fiber with regard to virion assembly, and particularly capsid assembly, is maintained by the presence of external amino acid trimerisation motifs such as the helical amino acid motif derived from the neck region of human lung surfactant protein D (Hoppe et al., FEBS Letters, 344: 191-195 (1994)). This and other trimerisation motifs known in the art can be functionally engineered into the fiber shaft and act as a fiber trimerisation signal to create knob-less fibers. For example, trimerisation motifs suitable for inclusion in modified viruses are described in WO 98/54346 and WO 99/41359. In a preferred embodiment of the invention, the non-native trimerisation motif present in the non-native polypeptide is the neck region peptide from human lung surfactant D.

As described in the Examples below, it may be convenient or necessary for any such additional or external motif or feature to be incorporated into the virus by means of a "linker" sequence. Such construction techniques for incorporation of DNA or amino acid sequences, via attachment to a linker sequence are known in the art and are within the routine skill of a protein/genetic engineer.

A "framework moiety" as defined herein, is a polypeptide (e.g. a protein) structure which retains a functional (e.g. folded) structure (or conformation) in the nuclear

and cytosolic cellular environments. Such functional structures are structures which allow a binding moiety attached to, or incorporated within, the framework moiety to retain ligand binding conformation in the absence of the ligand. This facilitates subsequent binding to that ligand, for instance, once the reassembled virion has left the cellular environment.

A framework moiety may thus be regarded as a type of "molecular scaffold" structure, which provides a framework to support or "hold" the binding moiety in an appropriate presentation or conformation to permit binding to its ligand. The framework moiety also provides the intramolecular interactions making a stable conformation in the cytosol possible. The framework moiety may thus be a protein or polypeptide molecule, which assumes a particular conformation or structure, and which tolerates modification in a particular region or regions, for example modification by amino acid sequence addition, insertion, deletion or substitution, or indeed insertion of a polypeptide sequence.

A "binding moiety" as defined herein, is thus a polypeptide structure attached to or comprising part of a framework moiety of the invention and which retains a binding function for a desired (target) ligand after expression of the non-native polypeptide of which it is part in the nucleus and cytosol of a host cell.

The binding moiety may be a contiguous or non-contiguous sequence of amino acids, and may be viewed as providing a binding site for the target ligand. A binding moiety may thus be provided by a region of a protein or polypeptide molecule, for example a surface region, e.g. a number of "surface" amino acid residues.

In certain preferred embodiments of the invention, the

framework and binding moieties may be of different origin, e.g. obtained from, or derived, from different sources, (e.g. from different proteins), for example by "grafting" or linking a desired binding moiety onto or to a desired framework. In other preferred embodiments, a binding moiety may be created within a protein "framework" structure, by modifying certain regions or residues of the framework protein, as described further below. In such an embodiment, the framework moiety may be represented by the "constant" or unmodified residues, and the binding moiety by the "variable" or modified residues.

A non-native polypeptide according to the present invention may thus be a combinatorial protein, that is a protein made by randomisation (random mutagenesis) of a particular protein structure, to generate a binding protein with novel, modified or enhanced binding characteristics. Such synthetically constructed "artificial" (in the sense of non-native) proteinaceous affinity binding molecules (i.e. proteins engineered to possess a particular or novel binding function) are known generally in the art.

Such combinatorial proteins can be prepared using various peptides and proteins as starting structures (Nygren and Uhlén, Current Opinion in Structural Biology, 7:463-469, 1997). Such proteins are known in the art, and may typically be prepared by random mutagenesis of a target protein, expression of the full library of these variants, e.g. on the surface of filamentous bacteriophage, followed by selection of a protein exhibiting the desired binding characteristics. This selection may typically involve a binding reaction between the variant protein and a target ligand (binding partner), which conveniently may be immobilised, and may be carried out *in vivo* or *in vitro*. The mutagenesis is

random in that the resulting amino acid encoded by any particular codon is not generally pre-determined but the positions where mutations are to be introduced are generally identified in advance. The mutagenesis may involve amino acid substitution, deletion, or addition (e.g. insertion).

Preferred combinatorial proteins for use as non-native polypeptides of the invention are proteins known generally in the art as affibodies. The term "affibody" as used herein defines an affinity binding molecule which is derived from a bacterial receptor protein, or binding domain thereof, wherein the binding domain is modified (e.g. by protein/genetic engineering) to modify (e.g. alter or enhance) the binding properties thereof. Advantageously, the affibody is a non-native protein (in the sense of not occurring in nature) and is further preferred to have a novel binding site. Examples and further descriptions of such protein molecules are given in WO 95/19374.

The use of an expression system such as surface display on phage provides a crucial link between genotype and phenotype; there is a self-contained unit which can be selected on the basis of its specific binding interactions and which also carries the nucleic acid encoding for the protein responsible for the observed binding characteristics. This enables expression in useful amounts of the protein selected for its binding characteristics, such expression typically taking place in a transformed bacterial host.

The protein, selected by its ability to bind to a target ligand (e.g. a desired cell surface molecule) may then be used to prepare the modified virus or viral component of the present invention, or more particularly a nucleotide sequence encoding the desired combinatorial

protein may be so used.

Techniques for construction of a combinatorial library of protein molecules and subsequent selection to obtain proteinaceous binding molecules having desired binding characteristics are known in the art (Nygren, P. and Uhlén, M. Current Opinion in Structural Biology (1997) 7: 463-469). Generally, a protein molecule, perhaps having intrinsic beneficial properties such as temperature or pH insensitivity, or conformational stability, is used as a "scaffold" or "framework" and a combinatorial library is then constructed via random but targeted amino acid substitutions (or other mutations) of that protein molecule, in order to produce a library of molecules having different binding characteristics. Surface residues are generally targeted for random mutagenesis.

In addition to phage display technology (Smith et al., Meth. Enzym. 217, 228-57, (1993)), other methods for library construction and selection include, for example, ribosomal display (Hanes et al., Proc. Natl., Acad. Sci. USA 94: 4937-4942 (1997)), peptides-on-plasmids (Schatz et al., Methods Enzymol., (1996) 267: 83-109), RNA-protein fusion (Roberts et al., Proc. Natl. Acad. Sci. USA 94: 12297-12302 (1997)) and DNA-protein linkage (STABLE) (Doi et al., FEBS Lett, 457(2): 227-30, (1999)).

Suitable protein frameworks may simply be linear peptides but preferably the framework will possess a folded three dimensional structure which has the potential for higher affinities and is less susceptible to proteolytic degradation. Although a framework may be designed *de novo*, naturally existing proteins or domains are usually selected for further engineering. For the avoidance of doubt, it is to be noted that throughout this specification the word "protein" is used to refer

to whole protein molecules as well as domains or fragments thereof, polypeptides or peptides.

The choice of protein framework depends on several parameters including an ability to be effectively expressed in a desired host cell (e.g a mammalian cell). The protein should also comprise sufficiently large regions on its surface which are tolerant to substitution (or insertion or deletion etc.) without losing the overall three dimensional structure. If the library is to be produced synthetically, a small overall size is a prerequisite. Where the selected framework protein has a binding function, amino acid residues involved in that interaction may be a target for randomisation. Randomisation may be performed in order to enhance known binding properties or to develop binding molecules with new specificities.

Suitable framework molecules are discussed in Nygren et al. (supra) and include cyclic peptides in a constrained sequence (the number of amino acid residues in such a constrained sequence is not critical and can be 5 or more, e.g. 5 to 10 or more, e.g. 40 or more), immunoglobulin-like scaffolds including Fv or single-chain (scFv) domains, bacterial receptors such as the 58-residue one-domain *Staphylococcal* protein A (SPA) analogue Z (the "Z Domain" being a derivative of the B domain of SPA), or other domains or analogues of SPA, DNA-binding proteins particularly zinc fingers and protease inhibitors.

Of particular interest is the bacterial receptor domain Z. Nord et al., in Protein Engineering 8(6):601-608 (1995), describe a method of constructing a combinatorial library of protein molecules based on the Z domain, which can be applied to a range of framework molecules. The method described is solid-phase-assisted

and based on the stepwise assembly of randomised single-stranded oligonucleotides.

As an alternative to modifying amino acid sequences or residues within a molecule to create a binding moiety, a binding moiety may be introduced to a framework moiety, for example by insertion or addition of a polypeptide constituting or comprising a binding moiety. The binding moiety may thus be attached to a framework moiety. In such a case, a binding moiety may be provided by an affinity binding partner for the desired target ligand.

Preferred binding moieties of the invention may be derived from, but are not limited to, ligands (i.e. binding partners) for cell surface receptors, anti-receptor antibodies (or antibody fragments or derivatives), cell specific peptides, single chain antibodies (ScFv), single domain antibodies, and minimal recognition units of antibodies such as a complementary-determining regions (CDRs) of Fv fragments. A binding moiety may thus be obtained or derived from the antigen-binding site or antigen binding or recognition/region(s) of an antibody, and such an antibody may be natural or synthetic.

Also envisaged within the scope of the invention are binding moieties derived from peptides or polypeptides in any form, including hormones, antibodies, T cell receptors, affibodies and ligands identified from various protein libraries.

As mentioned above, an important feature of the non-native polypeptide is that conformation is maintained in the cytoplasm and nucleus of a mammalian cell, such that the binding function of the binding moiety is retained. As further mentioned above, this is achieved by

providing a non-native polypeptide which maintains correct folding in the cytoplasm and nucleus of a mammalian cell. It has been found that such correct folding may be achieved using a non-native polypeptide which does not rely on disulphide bonding for conformation (i.e. which does not contain di-sulphide bonds). As will be discussed in more detail below, a further preferred feature of a non-native polypeptide according to the present invention is the presence of a α-helical structure.

WO 95/19374 describes several framework proteins, which share the beneficial feature of not relying on S-S bonds for their conformation. Such framework proteins are useful as framework moieties according to the invention. These include domains of bacterial receptors such as staphylococcal protein A (ααα type), protein G (IgG binding parts, ββαββ type), protein L (ββαββ type), and protein G (HSA binding parts, ααα type). Framework moieties according to the invention can also be derived from various bacterial receptors, for example, but not limited to those listed in the table 1:

Table 1: Examples of G+ bacterial receptors

Receptor[ligand][a]		Origin
Fc[IgG]receptor	type I	*Staphylococcus aureus*
	type II	*Staphylococcus pyogenes*[group A]
	type III	*Streptococcus* group C, G, L
	type IV	bovine group G streptococci
	type V	*Streptococcus zooepidemicus* [group C]
	type VI	*Streptococcus zooepidemicus* S212
Fibronectin receptor +		*S. aureus*, streptococci
M protein		*Streptococcus pyogenes* [group A]
Plasmin receptor		Streptococci group A
Collagen receptor		*S. aureus*, streptococci

Fibrinogen receptor	streptococci groups A, C, G
Protein L[6 light chains]	*Peptostreptococcus magnus*
Protein H[human IgG]	*Streptococcus pyogenes* [group A]
Protein B[human IgA, A1]	*Streptococcus agalactiae* [group B]
Protein Arp[human IgA]	streptococci group A
Serum albumin receptor	streptococci groups A, C, G

[a] Ligand is indicated when not obvious from receptor name

Framework moieties according to the invention are not restricted to structures from bacterial receptors. Also useful in the invention are other polypeptides comprising α-helical structures, often referred to as α-helical coiled coils, which are known from different sources (Cohen et al., Science 263:488-489, (1994); Harbury et al., Science 262:1401-1407, (1993)). The coil making up the framework in these peptides consists of repeats of amino acid sequences containing characteristic positioned hydrophobic residues. The structure of α-helical coiled coils is not dependent on intra- or intermolecular disulphide bridges for stability. Examples of α-helical coiled coils are the neck region peptide from human lung surfactant protein D, members or the spectrin superfamily, the leucine zippers and parts of the hemagglutinin in influenza virus.

Particularly preferred as sources of non-native polypeptides according to the present invention are members of the three-helix bundle family (e.g. as exemplified by the Z-domain of staphylococcal protein A).

Thus, in one embodiment, the non-native polypeptide according to the invention comprises a framework moiety which is based on the structure of a domain derived from a bacterial receptor. Preferred structures of the

framework moieties of the invention are derived from or based upon the alpha-alpha-alpha (ααα) -three-helix bundle or the beta-beta-alpha-beta-beta (ββαββ) structure classes. Also preferred are structures based upon or derived from the α-helical coiled coil family, particularly a member of any of the 2-, 3- or 4-helical coiled coil families.

In an embodiment of the invention, it is envisaged that the binding moiety is present within one or more of the helical bundles and/or one or more of the loops connecting these bundles.

In a further preferred embodiment of the invention, the framework moiety is based on the structure of a domain derived from staphylococcal protein A, streptococcal protein G or *Peptostreptococcus magnus* protein L.

In a further particularly preferred embodiment of the invention, the framework moiety is a derivative of the immunoglobulin binding Z-domain from staphylococcal protein A (Nord et al., supra).

In such embodiments the binding moiety may be created by combinatorial protein engineering as discussed above. In other embodiments, the domain or protein selected (e.g. the Z-domain) may be used in wild-type form as the non-native polypeptide of the invention.

As mentioned above, through combinatorial protein engineering, e.g. targeted to surface-located residues of the Z-domain, libraries may be constructed from which novel variants (i.e. novel binding molecules) termed Z-domain affibodies, may be selected by binding the desired target ligand (Hansson et al., Immunotechnology 4: 237-52 (1999)).

Preferred Z-domain-based non-native polypeptides may comprise the following amino acid sequences:

VDNKFNKEXXXAXXEIXXLPNLNXXQXXAFIXSLXDDPSQSANLLAEAKKLNDAQAPK [SEQ. ID. NO.: 45];

and

VDNKFNKEXXXAXXEIXXXXXXXXXXQXXAFIXSLXDXXXXSANLLAEAKKLNDAQAPK [SEQ. ID. NO.: 46],

where X is any amino acid.

In the above polypeptides, the conserved (i.e. specified) amino acid residues may be regarded as constituting the framework moiety, and the variable residues X, as together providing the binding moiety.

Furthermore, framework moieties according to the invention need not be dependent on alpha-helices for their stability, provided that the required conformation for ligand binding by the binding moiety is retained after expression of the non-native polypeptide in the nucleus and cytosol of a host cell. Such frameworks for example include frameworks derived from certain antibodies and their derivatives, known in the art, which do not require the presence of disulphide bridges to maintain structure and can therefore be expressed as part of a recombinant viral component in the nucleus or cytosol of a host cell, whilst retaining a functional conformation and functioning as binding moieties for a desired ligand. Other antibody structures within the scope of the invention include antibody structures in which structurally relevant cysteine residues have been replaced with alanine residues, and which retain the binding specificity of the antibody.

Additional framework moieties are also envisaged within the scope of the invention, for example, sequences derived from antibodies (e.g. monoclonal antibodies) and antibodies converted into a single-chain format allowing for the construction of genetic fusions to, or the modification of, viral components. In such a case where an antibody-derived framework moiety is used, then similarly as described above, a binding moiety may be created by protein/genetic engineering techniques (e.g. by modifying certain amino acid residues) or a binding moiety may be introduced (e.g. linked or grafted) to the antibody framework.

Antibody fragments conferring a desired binding activity can be used as binding moieties through the grafting of complementarity determining loops or regions (CDRs) of a an antibody with a desired tropism into a framework moiety capable of productive folding, i.e. capable of retaining the binding specificity(ies) of attached or incorporated binding moieties, in the cytoplasm and subsequent transport into the cell nucleus where the virus assembly takes place. In other words, a binding moiety may be a CDR of an antibody.

Further, according to the invention, certain antibody structure frameworks can be used as framework moieties, for example in conjunction with antibody-based or derived binding moieties e.g. specificity determining loops (CDRs), resulting in the directed construction of antibodies suitable for construction of re-targeted viral components.

An antibody framework may be used as a framework moiety (for example to receive CDR loop(s) as binding moiety(ies)) according to the present invention, provided that it is capable of productive folding in the cytoplasm and subsequent transport into the cell

nucleus. Single chain Fv fragments from monoclonal antibodies without disulfide bonds have been produced (Proba et al., J. Mol. Biol., 275: 245-53, (1998)). Such ScFv antibodies provided they meet the functional criterion above, may be used on framework moieties according to the present invention. The binding moiety may be the binding site of the ScFv antibody itself, or it may be further engineered as discussed above.

Intracellular selection of functional antibodies from a polyclonal repertoire has also been achieved (Gargano et al., FEBS Letters 414: 537-40, (1997)). This selection technology may be useful in the identification of antibody fragments suitable for genetic retargeting of adenoviruses, namely suitable for use as framework, and/or binding moieties according to the invention.

Other antibody frameworks which may be used, include those based on camel antibodies which are naturally devoid of light chains, or certain VH regions derived from conventional antibodies.

An example of an antibody framework according to the invention that is capable of productive folding in the cytoplasm, is a particular anti β-galactosidase single chain Fv fragment. This single chain variable chain fragment (VK, linker, VH) reactive with β-galactosidase and capable of being expressed in the cytoplasm has been previously described (Martineau P and Betton J-M: J. Mol. Biol. 292, 921-929 (1999). Thus, the invention also provides a modified virus comprising a framework moiety which comprises a sequence encoding this anti β-galactosidase single chain Fv fragment [SEQ ID. 47.] and functionally equivalent variants thereof.

The invention is not limited to framework moieties based upon antibody and affibody (e.g. receptor) structures,

but is extended to include other structures provided these are capable of retaining the binding specificity of attached or incorporated binding moieties when expressed in the cytosol/nucleus of infected cells.

Two or more non-native polypeptides, e.g. having different binding specificities, may be present in a modified virus of the invention, or a non-native polypeptide may incorporate two or more different binding moieties, or indeed two or more different framework/binding moiety constructs.

Accordingly, the invention provides a modified virus which comprises a first non-native polypeptide which binds a target cell and a second non-native polypeptide which binds a production cell or permissive cell.

Non-native polypeptides according to the invention may thus be present in recombinant viral components as comprising bi- or multi functional framework moieties (or framework/binding moiety constructs) constructed through genetic fusion between two or more different framework moieties (framework/binding moiety constructs). Use of such framework moieties (or framework/binding moiety constructs) can, for example confer infectivity of multiple cellular targets to the recombinant virus.

Also provided by the invention is a modified virus in which the non-native polypeptide according to the invention comprises a cleavage site positioned in a location that enables a binding moiety of the non-native polypeptide to be cleaved from the modified virus, for example preceding the binding moiety before the distal end of the fiber relative to the assembled virion. Examples of suitable cleavage sites are sites susceptible to a Factor Xa enzyme or a protease such as

the human rhinovirus 3C protease. If cleavage is carried out at the genetic level, for example, within a nucleic acid molecule comprising a part or all of a recombinant adenoviral genome, a cleavage site susceptible to an appropriate restriction enzyme or ribozyme may be used, for example in a particular cell population in which the virus is produced or targetted.

Non-native polypeptides according to the invention are capable of transport through the nuclear membrane. As is known from the art, this is an essential feature of viral components during replication in a host cell where expression of such components takes place in the cytosol, and assembly of such components takes place in the nucleus of the host cell e.g. in adenoviral replication.

The non-native polypeptide of the invention may be viewed as performing a dual role or function, firstly as providing a new binding domain (i.e. conferring altered tropism) and secondly as functioning as a viral component, namely a role as a functional part of a viral component (viral protein e.g. viral capsid protein). The non-native polypeptide may thus contribute to or function as a part of a viral protein. In other words, the non-native polypeptide may play a structural role or function, in the construction or assembly of a virus particle.

Cells comprising, or infected by, the modified virus of the invention also fall within the scope of the invention. Such cells may be of any origin provided that the cell type is capable of harbouring or propagating the virus. Generally, however, such cells will be mammalian cells. Such cells may be transfected with a membrane component (e.g. a membrane protein) that permits infection of the cell by the modified virus and

thus permits propagation of a virus which has a non-native tropism in a specific cell line, for example where propagation in conventional cell lines known in the art is prevented by way of the altered tropism of the virus.

Nucleic acid molecules encoding the non-native polypeptides, or modified viral component proteins or modified viruses of the invention are also envisaged. Vectors comprising such nucleic acid molecules, or the nucleotide sequence encoding the non-native polypeptides, or modified viral component proteins or modified virus of the invention, either for propagation of modified virus or further engineering of modified viruses are also included within the scope of the invention.

In a further aspect the invention provides a method for producing a modified virus according to the invention in cell culture. In one embodiment, the method comprises the steps of: genetically modifying a virus to produce a modified (e.g. recombinant) virus containing a non-native polypeptide, which polypeptide comprises one or more framework moieties each containing one or more binding moieties, which polypeptide is capable of being expressed in the cytoplasm and nucleus of a human host cell and there assuming and maintaining a conformation in the absence of a ligand for said binding moieties, which allow said binding moieties to subsequently bind with a said ligand and which polypeptide is capable of transport though the nuclear membrane, wherein said recombinant virus has an altered tropism conferred by said binding moieties; infecting permissive cells with such a virus; culturing said cells to produce the virus (e.g. at a sufficiently high titre), and, harvesting, and optionally, purifying the modified virus produced.

Ad vectors can be made replication competent or incompetent for permissive cells. For tumour therapy, replication competent Ad has the potential advantage that it can replicate and spread within the tumour (Miller et al., Gene Therapy, 3: 557-559). This may theoretically result in an increase of the chosen effector mechanism over that obtainable with replication incompetent vectors. Furthermore, infectious virus may contribute to an anti-tumour effect by cytopathogenic effects in infected cells as well as by evoking an anti-viral immune response which may harm infected cells.

It is desirable that a means of controlling the replication of competent modified virus is available in gene therapy applications.

Other means of controlling the replication of modified adenovirus are within the scope of the present invention For example, the modified virus according to the invention may comprise a gene encoding a viral protein required for viral replication present under the control of an inducible promoter or genetic element. For example, adenovirus pre-terminal protein (pTP), may be present under the regulation of a tetracycline responsive transcription activator (trTA) such that the pTP is only expressed in the presence of doxycycline. Alternatively, the modified virus according to the invention may comprise modifications to the genome such that the virus only replicates in cells which have a defect in the DNA synthesis - apoptosis regulatory pathways.

Conveniently this may be achieved in a modified virus of the invention by introducing the further feature of a cleavage site upstream of the binding moiety, or positioned in any such location that enables the binding moiety to be cleaved from the modified virus, e.g.

between the fiber shaft and the binding moiety.

Thus in a further aspect, the invention also provides a method of regulating the replication of a modified virus comprising the steps of: constructing a modified virus such that a cleavage site (e.g. a site susceptible to enzymatic or chemical cleavage) is positioned between a binding moiety required for cell infection and the remainder of the recombinant viral component of which the binding moiety forms part, and, bringing said recombinant virus into contact with a cleavage agent or cleavage means (e.g. an enzyme, chemical or in the case of a photo-labile cleavage site, light) capable of cleaving said binding moiety from said viral component and thereby preventing the recombinant virus from undergoing further infection cycles.

Where the means of cleavage in the method is enzymatic, the cleaving enzyme can be encoded within the genome of the recombinant virus and can be inducible. In a preferred embodiment the cleavage site is cleaved by a Factor Xa enzyme. The use of other proteases known in the art is also envisaged, for example, the human rhinovirus 3C protease, which is available as a fusion protein with GST (Walker et al., Bio/Technology 12:601 (1994)). This protease is active at 4°C, and recognises and cleaves the sequence LEVLFQ // GP.

Non-native polypeptides of the invention can be selected from libraries after screening of such libraries for correct nuclear and cytosolic folding of the peptide and a desired binding function in a manner similar to phage display techniques as known in the art. Such libraries may consist of candidate peptides fused to wild type or modified (e.g. recombinant) viral components according to the invention, for example adenoviral fiber proteins. A fiber fusion library can be expressed on adenovirions

and used to select for correct cytosolic and nuclear conformation by observing those candidate fusions which facilitate replication.

Modified virus according to the invention may also be constructed with both wild type and modified viral components present, or encoded in the genome, and if desired, with each component gene under the control of different genetic control elements, for example promoters. Thus, a modified virus according to the invention may comprise a modified viral component (i.e. modified to include or comprise a non-native polypeptide) and an equivalent or corresponding viral component which is unmodified, for example a wild-type fiber, and a modified fiber.

For example, a recombinant adenovirus can be constructed with a wild type fiber and a modified or recombinant fiber (i.e. a modified fiber comprising or incorporating a non-native polypeptide) e.g. a modified fiber derived from a 'fiber-candidate binding peptide' fusion library, which is expressed under control of a different promoter in an E1 deleted adenovirus. For example, the inventors have shown that a recombinant adenovirus fiber gene under the CMV promoter can be cloned into the multiple cloning site of the shuttle vector pAdTrack and that viruses can be produced that express both the WT fiber and the recombinant fiber by homologous recombination using the Ad vector pAdEasy (He T-C, Zhou S, DaCosta LT, Yu J, Kinzler KW and Vogelstein B: A simplified system for generating recombinant adenoviruses. Proc Natl Acad Sci, USA, 95:2509-2514, 1998).

It is also possible to have one or more fiber genes (for example, a wild type gene and a modified gene) present in an adenoviral vector genome under the control of one or more inducible promoters, thus allowing each gene to

be switched on or off independently. For example, in a desired Ad construct, the wild type fiber gene, displaced from its normal position and its control by the major late promoter (MLP), can be cloned under control of the hormone inducible MMTV promoter, inducible by dexamethasone. This promoter allows for expression of the wild type fibre in the vector propagating cells. preferably, the modified fiber is cloned in the same Ad vector genome, for example, downstream from a TRE (tetracyclin-responsive element) sequence element that is switched on the in the desired cell line. In cells expressing the tTA protein (a transcriptional activator which binds to TRE), TRE will be activated in the absence of Tc (tetracyclin) or Dox (doxacyclin). Alternatively and more favourably, Ad vector constructs can be propagated in cells which express reverse tTA (rtTA) where TRE, and expression the modified fiber, is activated in the presence of Tc (or Dox).

The recombinant fiber proteins in the library can have the wild-type fiber knob replaced by an external trimerisation motif and one or more members of a peptide library (or other non-native polypeptide). Screening of the recombinant fiber proteins for a desired binding activity can be carried out in a manner analogous to phage display once the recombinant adenovirus has been propagated via expression of the wild type fiber in a culture of propagator cells.

Accordingly, modified virus with an altered tropism will require cells which it is capable of infecting in order to allows its propagation. Thus the invention also provides permissive cells for virus according to the invention which are capable of being cultured to propagate the virus. A replication incompetent virus (which generally comprises a deletion in its genome

rendering it incompetent) needs special producer or propagator cells which are able to supply the genetic information that is deleted or missing from the virus, in order to replicate. Such cells or cell lines are known in the art. A modified virus with an altered tropism may be unable to infect such cells as are known in the art. Accordingly the invention includes the modification of such a propagator cell to include a binding partner (ligand) for the binding moieties of the modified virus (e.g. such that the cells express such a ligand on their surface).

Accordingly the invention provides a cell, preferably an *in vitro* or *ex-vivo* eukaryotic cell, containing a modified virus or viral component of the invention.

It is also envisaged that modified virus with an altered tropism might propagated in the same cells as the wild type virus by means of positioning the new tropism-altering binding moieties within a viral component which also comprises a wild type binding moiety (i.e. knob), with a cleavable site (e.g. for an enzyme such as Factor Xa or human rhinovirus 3C protease) between such moieties. After propagation in conventional cells via action of the wild type binding moiety, the viral component may be cleaved to remove the wild type binding moiety (e.g. the wild-type knob) and to reveal the new binding moiety and to confer the new tropism. It is also envisaged that the revealed binding moiety can be specific for a ligand which itself possesses a cell binding capability. For example, the binding moiety can be specific for the Fc region of an antibody, thereby permitting the use of an antibody which is cell specific for a desired cell type. Removal of the wild type binding moiety by cleavage and exposure of the virus to a cell specific antibody, allows the antibody to be adsorbed to the virus and thus targets the virus to the

cells specified by the antibody.

Non-native polypeptides according to the invention preferably confer an altered tropism to the recombinant virus in comparison with the tropism of the virus to which the non-native polypeptide has been introduced. Such altered tropisms allow recombinant viruses of the invention to be used in treatment of disease in human or animal subjects, either *by in vivo* treatment of, or *ex vivo* treatment of cells of, the subject requiring treatment.

The tropism of the modified virus of the invention may be altered such that the virus, via one or more binding moieties targets particular cells. Thus the invention also provides a modified virus wherein a non-native polypeptide comprises a binding moiety capable of binding to a cell specific ligand which may optionally be Prostate Specific Membrane Antigen, EGF receptor, Her-2/Neu, VEGF receptor, CD22, gp120, MHC/peptide complexes or membrane structures or surface molecules expressed or present on proliferating cells, tumor cells or virus infected cells

There are many ways in which modified virus exhibiting an altered tropism according to the invention can be used in gene therapy applications. In the case of tumour diseases, the following options exist:

I. Use of vectors to introduce transgenes into tumours, such as:

- antisense oncogenes
- suicide genes
- genes for immune modulatory substances or tumour antigens
- genes for anti-angiogenic factors

Examples of transgenes suitable for inclusion in the modified virus of the invention are antisense oncogenes, suicide genes, genes for immune modulatory substances, genes for tumour antigens, genes for anti-angiogenic factors, cytokines, genes for vascular endothelial growth inhibitors, genes for fusogenic membrane glycoproteins, cytotoxic genes, a gene encoding an enzyme which converts a pro-drug to cytotoxic substance, a gene for cytosine deaminase, a gene for uracil phosphoribosyl transferase. In a preferred embodiment of the invention the modifed virus comprises transgenes encoding cytosine deaminase and uracil phosphoribosyl transferase either as separate genes or, more preferably, together as a bifunctional fusion gene

In methods of treatment employing viruses of the invention with the transgenes cytosine deaminase and uracil phosphoribosyl transferase, preferably the virus is present or is co-administrated with one or more inhibitors of dihydropyrimidine dehydrogenase. This has the effect of increasing the toxic effect of the products formed by these transgenes from the pro-drug, 5-Fluorocytosine.

Also provided by the invention is a modified virus of comprising one or more viral components or transgenes present under the control of one or more inducible promoters or genetic elements, which are for example tissue specific or which respond to the presence of an exogenous expression modulating substance.

II. Use of infectious virus. This has the added value over the use of non replicating vectors in that virus can spread from cell to cell within a tumour, thereby multiplying the initial hit, or effect, on the tumour. Tumour cell destruction can occur not only by the cell-destroying mechanism engineered into the vector but

also by the cell destruction which is associated with the virus infection per se and by the attack of the body's immune response on the virus infected cells. This principle has already been tested in man through the direct intra-tumoural injection of an adenovirus which has been made gene manipulated to replicate only in p53 mutant tumour cells. The experience from these limited trials on large "head-and-neck" tumours are partially encouraging with a total regress of 2/11 treated tumours which are otherwise resistant to any form of known treatment (Shen Y: Personal communication).

Also provided by the invention is a modified virus wherein the gene encoding adenovirus death protein (ADP) is placed under the control of a promoter permitting over-expression of the protein, such that the lytic capacity of said modified virus is increased.

Thus, for gene therapy the modified virus may further be modified, according to standard techniques, principles and proposals widely described in the literature, to incorporate or comprise a desired therapeutic gene or therapeutic nucleic acid molecule. "Therapeutic" is used broadly herein to include both therapy (in the sense of curative or palliative therapy of a pre-existing or diagnosed condition) and prophylaxis. The gene may encode a desired therapeutic product (e.g. a therapeutic polypeptide or antisense molecule).

In a further aspect of the invention, the modified virus may include a site (e.g. a restriction site) for insertion of a desired therapeutic gene/nucleic acid molecule.

One disadvantage of using a wild type virus in gene therapy, even after modification according to the

invention, is that the human or animal subjects to be treated with the virus may possess pre-formed antibodies to components of the wild type virus still present or only partially modified in the modified virus. For example human adenovirus serotype 5 hexon protein is often targetted by pre-formed antibodies. Whilst modification of a single viral component, for example the component containing the non-native polypeptide according to the invention, may be sufficient to reduce the initial immune response to the modified virus via pre-formed antibodies, it may be preferable to modify a further component of the virus to further reduce the immune response of the host.

Thus the invention further provides a modified virus according to the invention which further comprises a viral component which is replaced with an equivalent component or is modified such that binding of said virus by antibodies pre-formed to the wild type virus is reduced. For example, an adenovirus of serotype Ad5 may comprise a hexon protein which is swapped for a hexon protein of a different serotype against which pre-formed antibodies are present at a reduced level compared to serotype Ad5, for example the different serotype may be Ad37. Alternatively, a modified adenovirus may comprise epitope sequences of the hexon protein to which pre-formed antibodies bind which are modified to produce a recombinant hexon lacking immunogenic epitopes.

Furthermore, the modified adenovirus may comprise a hexon protein which further comprises a peptide capable of binding a protein, for example, a non-immune system protein such as human serum albumen, sufficiently to cover the immunogenic epitopes of the hexon which are bound by pre-formed antibodies.

Also provided by the invention is a modified virus

according to the invention for use in therapy or in the preparation of a medicament for the treatment of tumour cells or proliferating cells.

Additionally provided by the invention is a pharmaceutical composition comprising a modified virus of the invention and a pharmaceutically acceptable carrier or excipient.

The invention will now be described in more detail with reference to the following non-limiting Examples, in which:

Figure 1 shows a schematic description of the sequence of construction of different recombinant fibers.
A. The NRP sequence is supplied with flanking Sph1 and Xho1 sites using PCR and ligated into WT fiber which has been supplied with flanking EcoR1 and Xho1 sites. The resulting fiber is called A1 and contains the fiber tail, first shaft repeat and the NRP motif.
B. EGF is joined to the fiber A1 by SOE. In the process an amino acid linker and a Cla1 restriction site is added between the NRP and EGF sequences. The resulting fiber is named A1 EGF. In this fiber, EGF can be substituted for new ligands by ligation using the Cla1 and Xho1 sites.
C. The NRP-Linker-EGF part of fiber A1 EGF is subjected to PCR. In the process an upstream Nhe1 site is introduced into the sequence in frame with the WT fiber sequence. After AT cloning, the sequence is ligated into WT fiber using Nhe1 and Xho1 to create fiber A7 EGF. A7 differs from A1 in that A7 contains the first seven shaft repeats of the Ad5 fiber. In the A7 EGF construct, EGF can be substituted for new ligands by ligation using the Cla1 and Xho1 sites.

Figure 2 shows a schematic representation of the

different recombinant fibers used in the present application.

A. The Ad5 wild type fiber.

B. The fiber A7 EGF described above.

C. The fiber A7 scFv C242 obtained by substituting EGF in fiber A7 EGF for the scFv fragment by ligation as mentioned above.

D. The fiber A7 scFv G250 obtained by substituting EGF in fiber A7 EGF for the scFv fragment by ligation as mentioned above.

E. The fiber A7 Affi IgG where EGF in fiber A7 EGF has been substituted for an IgG binding affibody as mentioned above.

F. The fiber A7 Affi IgA where EGF in fiber A7 EGF has been substituted for an IgA binding affibody as mentioned above.

G. G. The fiber A7 Affi IgG/Affi IgA where EGF in fiber A7 EGF has been substituted for an IgG binding affibody linked to an IgA binding affibody as mentioned above.

H. The fibre A7 ZIgG/ZIgG where EGF in fibre A7 EGF has been substituted for an IgG binding affibody linked to another IgG binding affibody as mentioned above.

I. The fibre A7 ZIgG Xa Knob where EGF in fibre A7 EGF has been substituted for an IgG binding affibody linked to a cleavable wild type fibre knob.

Figure 3 shows the binding of ZIgG affibody when incorporated into fibers expressed on virions.

Panel A. After binding of virus or protein to the membrane, it was incubated with Fc3(1) followed by HRP conjugated anti human IgG and developed. 1 = the virus A7 ZIgG Xa Knob; 2= the same virus after cleavage with Xa; 3 = WT virus; 4 = virus with two fibers e.g. WT and A7 ZIgG; 5 = Fc3(1); 6 = Protein A; 7 = Protein AG.

Panel B. After binding of virus or protein to the membrane, it was incubated with Fc3 followed by HRP

conjugated anti human IgG and developed. 1 = the virus A7 ZIgG Xa Knob; 2= the same virus after cleavage with Xa; 3 = WT virus; 4 = virus with two fibers e.g. WT and A7 ZIgG; 5 = Fc3; 6 = Protein A; 7 = Protein AG.

Panel C. Control to show presence of virus on membrane. 1 = the virus A7 ZIgG Xa Knob; 2= the same virus after cleavage with Xa; 3 = WT virus. After binding of virus the membrane was incubated with HRP conjugated anti Ad5 hexon and developed.

Panel D. Treated as membranes in panel A and B but the incubation with Fc was omitted. 1 = the virus A7 ZIgG Xa Knob; 2 = WT virus; 3 = protein A; 4 = protein AG.

Explanatory note: Fc3(1) is known to bind ZIgG and protein A. Fc3 is known to bind protein G (AG) only.

Figure 4 shows a schematic representation of the fibers described in Example 11. The NRP-linkers were inserted in the *Nhe*1 site upstream of NRP.

Figure 5 shows gel photos demonstrating expression and solubility of recombinant fibers in Sf9 cells. (A), whole cell lysate; (B), soluble fraction. Baculovirus-infected cells were harvested at 48 h after infection, lysed in isotonic buffer, and cell lysates divided in two aliquots. One aliquot was centrifuged at 10,000 x g for 10 min and supernatant was kept (panel B), whereas the other was analysed as whole cell lysate (panel A). Both aliquots were heat-denatured in SDS-sample buffer, analysed by conventional SDS-PAGE, and blotted. Blots were reacted with 4D2.5 mAb (specific for the fiber tail) and radiolabeled secondary antibody. Immunoblots were quantitated by autoradiogram scanning. Quantitative data, expressed as the percentage of soluble versus total fiber content, are shown in Table 7.

EXAMPLES

General procedures and starting materials.

Recombinant adenovirus fibers were constructed using methodologies based on ligation and PCR (Clackson et al., General application of PCR to gene cloning and manipulation, in PCR, A Practical Approach, Eds McPherson MJ, Quirke P and Taylor GR, IRL Press, Oxford, page 187, (1992)), i.e. PCR-ligation-PCR (Alvaro et al., BioTechniques 18: 746-750 (1995)) and splicing by overlap extension (SOE) (Horton et al., Recombination and mutagenesis of DNA sequences using PCR, in McPherson MJ (ed), Directed Mutagenesis, IRL Press 1991, p 217.). Gene products generated by PCR were generally cloned into the vector pCRII (Invitrogen Corp.) using so called TA cloning (Clark J.M., Nuc. Acids Res. 16: 9677-86, (1988)). Subclonings were performed according to standard methods (Sambrook et al., Molecular cloning. A laboratory manual. Second Edition. Cold Spring Harbor Laboratory Press, (1989)) in the vector pGEX-4T-3 (Amersham Pharmacia Biotech). Genes encoding recombinant fibers were sequenced using the Perkin Elmer ABI Prism sequencing equipment and were expressed in mammalian cells (SV40 transformed African Green monkey kidney cells, COS7, obtained from American Type Culture Collection, VA, USA) using vectors described below, and in insect cells (Sf9 cells from S. frugiperda obtained from American Type Culture Collection, VA, USA) using Baculovirus expression (Kitts et al., Biotechniques 14(5): 810-7, (1993)) (virus and vector from Clontech, Palo Alto, CA, USA) and stained with monoclonal antibodies specific for fiber tail, trimeric fiber and the new cell binding ligand. The following parameters were evaluated by immunostaining:

i) nuclear transportation

ii) functional expression of the new cell binding ligand

iii) ability to form trimers

Recombinant fibers were rescued into the Ad genome by a recently developed procedure (as described in Example 7 herein). The plasmid pTG3602 (Chartier et al., J. Virol., 70: 4805-4810, (1996)) containing the entire Ad5 genome as a Pac1-Pac1 fragment was used as starting material. The approximately 9kb fragment of the genome between Spe1 and Pac1 and containing the wild type fiber gene was cloned separately in pBluescript. From this fragment an approximately 3kb fragment between Sac1 and Kpn1 was further subcloned. A deletion of the native fiber gene with the exception of the N-terminal nucleotides upstream of the Nde1 site of the fiber was created in the 3kb fragment and an Xho1 site introduced in its stead allowing for ligation of recombinant fibers into the fiber-deleted 3kb fragment (the 3 kb fiber shuttle) between Nde1 and Xho1.

The 3 kb fiber shuttle with recombinant fiber was re-introduced into the 9 kb fragment cut with Nhe1 using homologous recombination in E.coli (Chartier et al., Supra). The resulting recombinant 9 kb fragment was finally excised from the vector with Spe1 and Pac1 and joined to the isolated 27 kb fragment by Cosmid cloning.

The presence of an insert of the expected properties was verified in all cosmid clones by PCR. Cosmid clones were also restricted with Hind III and the presence of restriction fragments of the expected size verified on gels.

Recombinant Ad genomes were isolated after restriction with Pac 1 and used to transfect suitable cells. The occurrence of plaques was determined by microscopic inspection of the transfected cell cultures.

List of oligonucleotide primer names for the primers used (all given as 5'-3' sequence):

[SEQ. ID. NO.: 20] Name 149: TTCCTCGAGTTATTCTTGGGCAATGTATGA

[SEQ. ID. NO.: 21] Name 175: GGGGAATTCGATGAAGCGCGCAAGACCGTCTGAA

[SEQ. ID. NO.: 22] Name: 196: GCTCGAGTTATCCGTTTGGAAACAACTCTAC

[SEQ. ID. NO.: 23] Name: 228: CTCGAGTCATCTCAATTCCCACCACTT

[SEQ. ID. NO.: 24] Name: 238: TGGCATGCCTGACGTAGCAAGCTTACGA

[SEQ. ID. NO.: 25] Name: 253:
GGGGAATTCATCGATGCAGGTCCAGTTGGTGCAGTCT

[SEQ. ID. NO.: 26] Name: 265: CAGGTCCAGTTGGTGCAGTCT

[SEQ. ID. NO.: 27] Name: 269:
GGGGGCCTGGGCGTCGTTCAGCTTCTTGGCTCCGTTTGGAAACAACTCTAC

[SEQ. ID. NO.: 28] Name: 270:
CTGAACGACGCCCAGGCCCCCAAGAGCGACCCATCGATCATGAACTCCGACTCCGAATGT

[SEQ. ID. NO.: 29] Name: 273:
CCCCTGGAGTTAAATTTTCTTGTCCACCTTGGTGCT

[SEQ. ID. NO.: 30] Name: 274:
GGGGAATTCATCGATGGACTACAAAGATATTGTGATGACGCAGGCT

[SEQ. ID. NO.: 31] Name: 275: CTACCTCGAGTTAACACTCATTCCTGTTGAAGC

[SEQ. ID. NO.: 32] Name: 326: GGGGCTAGCCCCTGACGTAGCAAGCTTACGA

[SEQ. ID. NO.: 33] Name: 403: GGG CTC GAG TTA CTC GAT GGG GGC TGG GAG GGC

[SEQ. ID. NO.: 34] Name: 414:
GGCCCCCGAGGCCTCGAGTGAGGAGACGGTGACCGTGGT

[SEQ. ID. NO.: 35] Name: 416:
GGCCCAGCCCACGAATTCATCGATGGATATTGTGATGACGCAGGCT

[SEQ. ID. NO.: 36] Name: 418: AGA CTG CAC CAA CTG GAC CTG $(SNN)_{18}$CCGTTTCAGCTCCAGCTTGGT (S is dA, dG or dC and N is dA, dG, dC or dT)

[SEQ. ID. NO.: 37] Name: 473: GGC AAT TCC ATC GAT CGC CAC CAT GGA CAT TGT GAT GAC CCA GTC T

[SEQ. ID. NO.: 38] Name: 474: CCC CTC GAG TTA ACA CTC ATT CCT GTT GAA GCT

[SEQ. ID. NO.: 39] Name: 476: ACC ACG GTC ACC GTC TCC TCA GCT GAT GCT GCA CCA ACT GTA

[SEQ. ID. NO.: 40] Name: 478: TGA GGA GAC GGT GAC CGT GGT

[SEQ. ID. NO.: 41] Name: 503: GGGCCATCGATCGTAGACAACAAATTCAACAAA

[SEQ. ID. NO.: 42] Name: 504: GGGCTCGAGTTATTTCGGCGCCTGAGCATCATT

[SEQ. ID. NO.: 43] Name: 550:

TCGGTTTGGAAACAACTCTACCTTTTTTTTCGGCGCCTGAGCATCATT

[SEQ. ID. NO.: 44] Name: 551:

AAAAAGGTAGAGTTGTTTCCAAACGGAGTAGACAACAAATTCAACAAA

EXAMPLE 1:

Genetic insertion of a trimerisation motif (the neck region peptide from human lung surfactant D) into adenovirus fibers.

The gene encoding Ad 5 WT fiber was obtained from a preparation of Ad5 virus by PCR using an upstream primer (Primer 175) identical to the first six coding triplets of the fiber plus an EcoR1 site and a downstream primer (Primer 149) annealing to the six terminal coding triplets of the fiber plus an Xho1 site. The fiber thus obtained [SEQ. ID. NO.: 1] was cloned into the vector pBluescript using these restriction sites and can be further sub-cloned into other vectors using the same restriction enzymes.

Fiber peptides were made where the knob was replaced with an external trimerisation motif (see below). The purpose behind the introduction of an external trimerisation motif is two-fold: a) to remove the knob containing the native trimerisation signal but also the cell binding part of the fiber, and b) simultaneously to supply the necessary trimerisation signal. In this case one particular amino acid motif have been used, i.e. the 36 aa "Neck Region Peptide" = NRP [SEQ. ID. NO.: 2] from human "Lung Surfactant Protein D" (Hoppe et al., Supra). It should be noted that the sequence used is slightly longer than the actual trimerisation part of NRP in that the eight amino acids (KKVELFPN) following the trimerisation signal in human lung surfactant protein D

has been retained in all constructs containing the NRP sequence. The sequence KKVELFPN functions as an efficient linker between the trimerisation signal and the C-terminal carbohydrate domain of the human lung surfactant D and is considered to have the same important function in the recombinant fibers described herein. The DNA sequence coding for the trimerisation motif was synthesized, cloned and verified by sequencing.

In order to introduce the NRP motif into the adenovirus fiber, the NRP sequence was subjected to PCR (Clackson et al., Supra) with the upstream primer 238 containing Sph1 N-terminally of the NRP coding sequence and the downstream reverse primer 196 containing Xho1 C-terminally of the coding sequence. After cutting with Sph1 and Xho1 the NRP sequence was ligated into the WT fiber gene cut with the same enzyme. The resulting recombinant fiber A1 [SEQ. ID. NO.: 3] contains the fiber tail and the first shaft repeat followed by the NRP trimerisation motif. For a schematic representation of the constructions and construction pathways see Figs 1 and 2.

To replace the cell binding function of the knob a new cell binding ligand was subsequently introduced into the fiber in addition to the external trimerisation amino acid motif (see below).

EXAMPLE 2:

Assembly of gene construct encoding recombinant adenovirus fibers with epidermal growth factor (EGF) and the external trimerisation motif from human lung surfactant D.

For a schematic representation of the constructions and construction pathways see Figs 1 and 2.

The DNA sequence for human EGF [SEQ. ID. NO.: 4] was synthesized, cloned and sequenced in the project. This sequence was then joined to the A1 fiber mentioned above by splicing by overlap extension. In this case the EGF gene was subjected to PCR with an upstream primer (270) (identical to the first seven coding triplets) and containing an overhang with the sequence for an amino acid linker [SEQ. ID. NO.: 5] derived from Staphylococcal protein A and a Cla1 restriction site, and a downstream primer (228) (complementary to the seven terminal triplets of the + strand and containing an Xho1 site). The A1 fiber gene was subjected to PCR with an upstream primer identical to the first six coding triplets of the gene and an EcoR1 site (175) and a downstream primer (269) complementary to the seven terminal coding triplets of the + strand and an overhang complementary to the overhang in Back primer for the EGF sequence. The two PCR products where then joined by PCR under standard SOE conditions (Horton et al., Supra) to produce fiber A1 EGF [SEQ. ID. NO.: 6].

In order to construct a fiber with the first seven shaft repeats, the NRP trimerisation signal, the Staphylococcal linker and EGF, the fiber A1 EGF was subjected to PCR with an upstream primer (326) identical to the first seven 5' triplets of the NRP sequence plus an upstream Nhe1 site and a downstream primer (228) complementary to the seven terminal triplets of the A1 plus strand. After cloning the PCR product was restricted with Nhe1 and Xho1 and ligated into WT Fiber restricted with the same enzymes to obtain Fiber A7 EGF [SEQ. ID. NO.: 7] which is similar to A1 EGF but differs in that it contains the first seven shaft repeats of the Ad5 Fiber.

Example 3:

Assembly of gene construct encoding recombinant

adenovirus fibers with single-chain antibodies and the external trimerisation motif from human lung surfactant D.

Two monoclonal antibody single chain fragments were used to construct recombinant adenovirus fibers. The first is a single chain fragment (scFv) of the monoclonal antibody G250, which with high selectivity has been shown to react with a protein antigen on human renal carcinoma cells (Oosterwijk et al., Int. J. Cancer 38: 489-94, (1986)). The second is a single chain fragment of the monoclonal antibody C242 which reacts with i.a. colorectal and pancreatic carcinomas (Johansson C., Thesis, University of G`teburg, (1991)).

G250 constructs

The single chain fragment (Variable kappa chain or VK, linker, variable heavy chain or VH, joining sequence and constant heavy domain 2 or CH2) of the antibody G250 was constructed as previously described (Weijtens et al., J.Immunol., 152(2): 836-43, (1996)). This G250 construct is [SEQ. ID. NO.: 8]. To permit cloning into the aforementioned A1 and A7 fiber constructs, the single chain fragment was supplied with an upstream ClaI site and a downstream XhoI site by PCR using primers 416 and 403.

C242 constructs

The single chain fragment of the antibody C242 (Variable kappa, Linker, variable heavy and constant kappa or CK) [SEQ. ID. NO.: 9] was constructed as follows by SOE using cDNA from the antibody producing hybridoma as original templates. VKCK was amplified using primers 274 and 275, VHCH1 was amplified using primers 253 and 273. An scFv (single chain variable fragment) (VK LinkLib VH) was constructed by SOE as follows. VK and VH were amplified separately using primers 416/418 and

265/414 respectively with the above mentioned VKCK and VHCH1 as templates and joined together by SOE using primers 414 and 416. In the construct VK LinkLib VH, the linker between VK and VH is a randomized 18 amino acid sequence as described previously (Tang et al., J. Biol. Chem., 271: 15682-86, (1996)). The nucleotide sequence for this linker is present in primer 418. The construct VK LinkLib VH was cloned into the vector pAK100 (Krebber et al., J.Immunological Methods 201: 35-55, (1997)). Phage display and selection of antigen binders by panning was performed using methods described earlier (Krebber et al., Supra). In the present experiments, the CanAg antigen, reacting with antibody C242, was adsorbed onto biotinylated antibody C241 (which binds another epitope on the antigen than C242) bound to streptavidin coated tubes (CanAg Diagnostics Ltd, G`teburg, Sweden). Several binders were isolated and shown to contain different linkers. A particular VK Linker VH construct shown by sequencing to contain the linker PPDFVPPAASFPDHSPRG (one letter amino acid code) was selected for further work based on antigen binding ability. CK was linked to this construct by SOE to obtain the format C242 VK Link VHCK. In this SOE the VK LinkLib VH was amplified using primers 416 and 478 and the CK amplified with primers 476 and 474. The amplified products were then joined by SOE using primers 416 and 474.

In the PCR reactions mentioned above the gene sequences for the single chain fragments were supplied with an upstream Cla1 site (present in primers 416 and 473) and a downstream Xho1 (present in primers 403 and 474) to allow for ligation into the A1 and A7 fiber constructs mentioned earlier (for a schematic representation of the constructions and construction pathways see Figs 1 and 2) to construct the fibers A1 G250, A7 G250, A1 C242 and A7 C242 (Fig 1 and [SEQ I.D. NOS. 10-13]).

Example 4:

Assembly of gene constructs encoding recombinant adenovirus fibers with affibodies and the external trimerisation motif from human lung surfactant D.

To investigate if mono and divalent antigen binding structures based on the structure of single staphylococcal protein A domains could be functionally expressed when incorporated into recombinant Adenovirus fiber, gene constructs for mammalian and insect cell expression were made. Assembly of gene constructs encoding recombinant adenovirus fibers containing the IgG binding Z domain (ZIgG) derived from staphylococcal protein A (Nilsson B., Prot. Eng. 1:, 107-13, (1987)) [SEQ. ID. NO.: 14] or a Z domain-derived IgA-specific affibody (ZIgA) selected using phage display (Gunneriusson E. et al., App. Env. Micro., 65: 4134-40, (1999)) [SEQ. ID. NO.: 15] was accomplished as follows.

The genes encoding the respective affinity moieties were amplified by PCR using primers 503 and 504 on the following plasmid templates; pEZZmp18 (Tang et al., Supra) for the Z domain construct and pKN1-dZIgA (Clackson et al., Supra) for the ZIgA construct.

In the PCR amplification, the genes for the two different affinity ligands were supplied with an upstream Cla I site and a downstream Xho I site in the appropriate reading frame for subsequent ligation into the above described fiber gene A7 EGF resulting in constructs encoding the recombinant fiber A7 ZIgG [SEQ. ID. NO.: 16] and A7 ZIgA [SEQ. ID. NO.: 17], respectively, which were adapted for later being rescued into the Ad genome (see below) for the production of recombinant viruses carrying the new binding specificities. Furthermore, the fiber A1 ZIgG was constructed by ligation of the modified ZIgG into the

aforementioned fiber A1 EGF (see also Fig 1).

In addition, a third gene construct was assembled encoding a fiber containing both of the above mentioned affinity domains. This construct encodes a fiber containing the fiber tail, the first seven shaft repeats, the NRP sequence, the staphylococcal protein A linker, the IgG binding Z domain, the eight amino acid linker from NRP (KKVELFPN) followed by the IgA binding affibody. To construct this fiber the two different affinity domains were first genetically joined together by SOE using primers 550 and 551 with overhangs complementary to the nucleotide sequence encoding the linker sequence KKVELFPN. In the PCR process an upstream ClaI site and a downstream XhoI site were introduced by the primers allowing for ligation into the vector pGEX-4T-3 containing the Fiber A7 EGF gene construct to obtain Fiber A7 ZIgG/ZIgA [SEQ. ID. NO.: 18].

A further gene construct was also assembled to encode a fiber containing two linked ZIgG domains. This gene codes for the fiber A7 ZIgG /ZIgG [SEQ. ID. NO.: 19] and was assembled exactly as described for A7 ZIgG /ZIgA with the exception that the gene for ZIgG was used instead of ZIgA in the PCR reaction.

For a schematic representation of the constructions and construction pathways see Figs 1 and 2.

EXAMPLE 5:
Binding studies

The genes encoding recombinant fibers were cloned into the vectors pcDNA (which targets proteins for expression in the cytosol) and pSecTag (which targets proteins for expression as secreted products), both from Invitrogen

BV, Groningen, The Netherlands, and transfected into COS7 cells using Lipofectamin (Life Technologies Inc, Gaithersburg, MD, USA) as described by the manufacturers.

Expression and cellular localization of recombinant fibers were evaluated by immunostaining using the following primary reagents:

Mouse monoclonal antibody 4D2.5 (anti-Ad5 fiber) (kindly provided by Dr Geoffrey Engler, University of Birmingham, Alabama, USA) (Shin Hong et al., Virology 185: 758-767, (1991)).
Mouse monoclonal antibody 2A6.36 (anti-trimerised Ad5 fiber) (kindly provided by Dr Geoffrey Engler, University of Birmingham, Alabama, USA) (Shin Hong et al., Supra).
Mouse monoclonal antibody against Epidermal Growth Factor (EGF) = a-EGF (Cambio, Cambridge, UK, Cat no CA 954).
Biotinylated mouse monoclonal anti idiotypic antibody directed against monoclonal antibody C242 = a-Id C242 (Lindholm et al., unpublished results).
Biotinylated mouse monoclonal anti idiotypic antibody directed against monoclonal antibody G250 = a-Id C250 (Kindly supplied by Reinder Bolhuis, Daniel Den Hoed Cancer Center, Rotterdam, The Netherlands).
Human polyclonal IgG and IgA for evaluation of Affibody activity = HIgG (Sigma-Aldrich Fine Chemicals, Cat no I4506) and HIgA (Sigma-Aldrich Fine Chemicals, Cat no I1010).

Secondary reagents were:
For identification of mouse antibodies: FITC labelled F(ab)2 rabbit anti mouse immunoglobulin (DAKO A/S, Glostrup, Denmark, Cat no F0313) = aMIg.
For identification of human IgG: FITC labeled F(ab)2

rabbit anti human IgG (DAKO A/S, Glostrup, Denmark, Cat no F0315) = aHIgG.
For identification of human IgA: FITC labeled F(ab)2 rabbit anti human IgA (DAKO A/S, Glostrup, Denmark, Cat no F0316) = aHIgA.
For identification of biotinylated antibodies: FITC labeled Streptavidin (DAKO A/S, Glostrup, Denmark, Cat no F0422).

Briefly, cells were centrifuged onto microscope slides in a Shandon Cytospin2 cytocentrifuge and air-dried over night at room temperature. The preparations were fixed in 3% paraformaldehyde in phosphate buffered saline, pH 7,4 (PBS), permeabilized with 0,1% Triton-X100 in PBS. After washing in PBS, preparations were incubated with primary reagents for 30 minutes at 37°C in a humid chamber, washed again in PBS and incubated with secondary reagent for 30 minutes at 37°C in a humid chamber. After washing in PBS, preparations were mounted in PBS with 50% glycerol and viewed in a Zeiss Axophot microscope equipped with appropriate light source and filters for FITC.

The results are shown below in Tables 2 and 3. It is obvious that all of the different ligands show appropriate binding when the corresponding fibers were expressed as secreted products. However, only the affibodies show the expected correct binding when the fibers were expressed in the cytosol and subsequently transported to the nucleus. Therefore, not all ligands can fold correctly in the cytosol and nucleus. It is interesting that the ligands which can withstand the milieu in the cytosol and nucleus (affibodies) are small α-helical structures not depending on S-S bridges for their conformation whereas the ligands which were shown not to be properly expressed in the cytosol/nucleus all have a conformation that is dependent on the formation

of S-S bridges which are formed very poorly or not at all in the cytosol.

Table 2: Expression and functional ligand binding in recombinant adenovirus fibers targeted for secretion in COS cells

	Detecting reagent						
Fiber	4D2+ aMIg	2A6+ aMIg	a-EGF+ aMIg	a-MIg	a-Id+ aMIg	HigG+ aHIgG	HigA+ aHIgA
WT	+	+	ND	ND	ND	ND	ND
A7 EGF	+	+	+	ND	ND	ND	ND
A7 G250	+	+	ND	+	+	ND	ND
A7 C242	+	+	ND	+	+	ND	ND
A1 ZIgG	+	ND	ND	ND	ND	+	-
A7 ZIgG	+	ND	ND	ND	ND	+	-
A7 ZIgG/ZIgG	+	ND	ND	ND	ND	+	-
A7 ZIgA	+	ND	ND	ND	ND	-	+
A7 ZIgG/ZIgA	+	ND	ND	ND	ND	+	+

Table 3:Expression and functionality of ligand binding in the nucleus of COS cells in native and selected recombinant fibers after targeting for expression in the cytosol

	Detecting reagent						
Fiber	4D2+ aMIg	2A6+ aMIg	a-EGF+ aMIg	a-MIg	a-Id+ aMIga	HigG+ HIgG	HigA+ aHIgA
WT	+	+	ND	ND	ND	ND	ND
A7 EGF	+	+	-	ND	ND	ND	ND
A7 G250	+	+	ND	+	-	ND	ND
A7 C242	+	+	ND	+	-	ND	ND
A1 ZIgG	+	ND	ND	ND	ND	+	-
A7 ZIgG	+	+	ND	ND	ND	+	-
A7 ZIgG /ZIgG	+	ND	ND	ND	ND	+	-
A7 ZIgA	+	+	ND	ND	ND	-	+
A7 ZIgG/ZIgA	+	ND	ND	ND	ND	+	+
A7 ZIgG Xa Knob	+	ND	ND	ND	ND	+	-

The results have obvious implications for the construction of those re-targeted virus for human gene therapy where the viral structural components containing the new cell binding ligand are synthesized in the mammalian cell cytosol, i.e. adenovirus. Below are two enabling examples to show how such re-targeted adenovirus can be constructed.

EXAMPLE 6:

Grafting of CDR loops

Certain single chain constructs of monoclonal antibodies retain their binding specificity even in the mammalian cell cytosol (Cattaneo et al., TIBTECH 17: 115-121, (1999)). This is a function of the so called frame work regions of the antibody variable regions. It is known

that the antigen binding CDR loops can be transferred from one antibody to another by recombinant DNA technology thereby creating antibodies with frame work properties from one antibody and binding properties from another (for methodology see Emery et al., Strategies for Humanizing Antibodies, in Antibody Engineering, Carl A.K. Borrebaeck, (ed.), Oxford University Press 1995, page 159).

Such a loop-grafted single chain antibody, based on a variable domain framework capable of folding in the cytosol and subsequent transport to the cell nucleus thus created can subsequently be supplied with appropriate cloning sites at the gene level, ligated into fiber A7 RGD encoding genes and rescued into the adenovirus genome as described below (Example 7).

EXAMPLE 7:

Rescuing of recombinant fibers into the adenovirus genome

The wild type fiber as encoded in the Ad genome was substituted for recombinant fibers by the following method developed within the project. In the method the wild type Ad5 genome in the plasmid pTG3602 (Emery et al., Supra) was used as receptor for genes encoding the recombinant fibers. This plasmid contains the entire wild-type Ad5 genome joined to the plasmid backbone by Pac1 linkers. The entire genome can be recovered as a linear DNA fragment after cleavage with Pac1 since Pac1 sites are absent from the Ad genome. The resulting linear Ad DNA can then be transfected to susceptible cells to yield virus (Chartier et al., Supra). From this plasmid the Ad genome can also be cleaved as two fragments, one of 27 kb and one of 9 kb, using the enzymes Pac1 and Spe1. The 9 kb fragment has been cloned into pBluescript. From the 9 kb fragment, which

contains the fiber, a 3 kb Sac1-Kpn1 fragment containing the fiber gene was further subcloned. The fiber gene was deleted between the Nde1 site in the tail-portion of the fiber and the Mun1 site which is situated just down-stream of the fiber-gene and an adapter containing an Xho1 site and the down-stream sequence was introduced between the Nde1 and Mun1 site to obtain a fiber shuttle vector. Several recombinant fibers have now been ligated between Nde1 and Xho1 of this shuttle vector and thereafter rescued into the 9 kb fragment mentioned above by homologous recombination in *E. coli* (Chartier et al., Supra). This means that all normal elements regulating fiber-expression have been left intact.

Finally, recombinant 9 kb fragments were separately joined to the 27 kb fragment by cosmid cloning to re-create the complete Ad genome.

The 27 kb fragment may also be derived from another adenovirus genome, such as the Ad-YFG described by He et al (He T-C, Zhou S, DaCosta LT, Yu J, Kinzler KW and Vogelstein B: A simplified system for generating recombinant adenoviruses. Proc Natl Acad Sci, USA, 95:2509-2514, 1998). If the 27 kb fragment is derived from pTG3602 the resulting genome will be WT E1+ whereas the 27 kb fragment from Ad-YFG will render E1 deleted viruses needing e.g. low passage 293 for replication.

The recombinant Ad genome resulting from above manipulations can finally be obtained as a linear fragment by cleaving with Pac1 and used to transfect permissive cell lines which then yield virus plaques if the genome is functional. For these transfections the FuGENE 6 transfection agent (Roche) was used.

Various recombinant fibers were rescued into the Ad genome and subsequently transfected into permissive

cells. Results are shown below in table 4. Of the fibers accounted for, only those containing the WT knob and the affibody A7 ZIgG/IgG were capable of rendering functional virions. The results are in complete concordance with those shown in tables 2 and 3.

Table 4: Rescuing of recombinant fibers into virions

Fiber	*Transfected cells*	*Occurrence of virus plaques*
A7 WT Knob	293	Yes
A7 EGF	A549	No
A7 C242	Colo 205	No
A7 G250	A75	No
A7 ZIgG/ZIgG	293/Fc*	Yes

(*) 293/Fc cells are 293 cells stably transfected with Fc3(1) from human IgG expressed as a membrane protein. To create 293/Fc, cloned Fc3(1) from human IgG which reacts with the Z-domain (Jendeberg, L., Nilsson, P., Larsson, A., Nilsson, B., Uhlén, M. and Nygren, P.-Å (1997) "Engineering of Fc1 and Fc3 from human immunoglobulin G to analyse subclass specificity for staphylococcal protein A" J. Imm. Methods 201, 25-34.) was ligated into the vector pDisplay (Invitrogen) in frame with the PDGFR transmembrane domain sequence present in this vector. The coding sequence containing the Fc sequence fused to the PDGFR transmembrane domain sequence was then cleaved from the vector using the restriction enzymes Sfi1 and Not1 and ligated into the vector pSecTag (which carries the Zeocin resistance gene) cleaved with the same enzymes. The reconstituted vector was the transfected into low passage 293 cells using FuGENE and the cells were placed under selection pressure using Zeocin to select for stably transformed cells. Clones were isolated and tested for membrane expression of membrane-bound Fc3(1) using FITC labeled

staphylococcal protein A (Sigma). One 293 clone which homogenously expresses membrane bound Fc was used for transfection of the A7 ZIgG/ZIgG containing genome to produce virions.

EXAMPLE 8:

Assembly of gene constructs encoding recombinant adenovirus fibers with an affibody and a cleavable wild type knob

Cleavable fibers containing both a non-native polypeptide comprising an external ligand, here an affibody, and a C-terminally placed wild type knob were constructed with an activated factor X site situated between the cell binding structures so that the knob can be cleaved off to expose the affibody. This permits virus production in 293 cells with subsequent infection of new target cells as defined by the affibody after proteolytic removal of the knob.

To construct a gene for a cleavable fiber for ligation of different ligands the WT fiber was subjected to PCR with an upstream primer (437) introducing Cla1, Mun1 and the Factor Xa recognition site before the first seven triplets of repeat 22 and the downstream primer 149 which primes at the end of the fiber knob. After cloning and restriction with Cla1 and Xho1 the DNA fragment was cloned into the A7 fiber construct mentioned earlier. The resulting fiber gene contains from the N-terminus the sequence for the tail, the first seven shaft repeats, the NRP trimerisation signal, the linker from staphylococcal protein A, a Cla1 site, a Muni site, the Xa cleavage site, repeat 22 and the wild type knob. For ligation into this fiber gene the gene for the affibody ZIgG was supplied with an upstream Cla1 site and a downstream Mun1 site by PCR using primers 503 and 505 and ligated into the aforementioned "Xa fiber"

gene. The resulting fiber gene contains from the N-terminus the sequence for the tail, the first seven shaft repeats, the NRP trimerisation signal, the linker from staphylococeal protein A, a ClaI site, the new linker, a MunI site, the Xa cleavage site, repeat 22 and the wild type knob.

The fiber gene was rescued into the adenovirus genome (see Example 7). The recombinant genome was transfected into 293 cells and virions were produced and purified on CsCl gradients. Purified virions were cleaved with activated Factor X (Xa) (Sigma) in 50mM Tris-Cl pH 8,0; 100 mM NaCl; 5mM $CaCl_2$ for 48 hours at room temperature. 1 U of Xa was used for 5µL of virus suspension.

Assay for binding of the ZIgG affibody on the virions was performed essentially as follows. Virus suspensions were blotted onto PVDF membranes (BioRad). After blocking with 3% gelatin in Tris buffered saline, pH 7,4 (TBS) the membranes were incubated with FC 10µg/mL in 1% gelatin in TBS for 90 minutes at room temperature, washed, incubated with anti-human IgG-HRP (Dakopatts) in TBS/1% gelatin for 90 minutes at room temperature, washed and developed in 4-Chloro Naphtol reagent (BioRad). The results of staining of WT and recombinant viruses are shown in Figure 3 along with different controls. It is obvious that the recombinant virus binds Fc3(1) which binds to WT Z whereas Fc3 which does not bind to WT Z (Jendeberg, L., Nilsson, P., Larsson, A., Nilsson, B., Uhlén, M. and Nygren, P.-Å. (1997) "Engineering of Fc1 and Fc3 from human immunoglobulin G to analyse subclass specificity for staphylococcal protein A" J. Imm. Methods 201, 25-34) fails to bind to the recombinant virus. WT virus does not bind either Fc preparation.

Example 9:

Assay of non-native polypeptides for functional conformation and binding specificity

The candidate polypeptide is expressed as part of a viral component protein by means of insertion of the corresponding coding sequence into a suitable construct. The construct, which comprises an entire viral genome in which one or more components have been replaced by the corresponding recombinant component gene, is expressed in a cell line suitable for propagation of the recombinant virus coded for by the recombinant viral genome. There are two main possibilities depending on whether the WT cell-binding function is retained in the recombinant genome or not.

If the WT cell-binding function is retained virus can be produced in normal producer lines such as 293 for adenovirus. The ability of the non-native polypeptide to be expressed in a conformation that allows it to bind to an extracellular ligand and to be part of a functional virion is assayed by binding studies on virions. This approach is exemplified by the strategy employed in Example 8 above where the virus fiber contains both the WT knob and a non-native binding ligand and where the knob can be removed by proteolytic cleavage. The binding function of the non-native polypeptide was demonstrated by standard solid phase immunotechnology. Another strategy to meet the same ends would be to produce virions with WT fibers as well as recombinant fibers, the latter containing the non-native polypeptide as replacement for the native knob.

If the WT cell-binding function has been deleted from the viral genome the virions will depend on the non-native polypeptide for cell binding. Therefore, the

cell line must be able to permit entry of the recombinant virus and, if necessary, to supply genetic information that may be missing from the recombinant virus, e.g. the sequence for E1. If the recombinant genome is E1 deleted, one solution is to use an E1 transfected cell line such as 293 and to stably transfect these with a receptor capable of binding the non-native polypeptide coded for by the recombinant viral genome. The ability of the non-native polypeptide to be expressed in a conformation that allows it to bind to the corresponding receptor structure and form part of a functional virion is assayed by screening for plaque formation in the cells following transfection of the viral genome. This strategy was employed in Example 7 above where 293 cells stably expressing membrane bound Fc from human IgG was used to produce virions where the WT binding function was replaced with the affibody ZIgG. The exact binding properties of the non-native polypeptide can further be determined in studies where the virus is allowed to compete for binding with peptides structurally related to the non-native polypeptide responsible for cell-binding.

EXAMPLE 10:

Anti β-galactosidase single chain Fv fragment

In a continuation of the experiments described in Example 3, a single chain variable chain fragment (VK, linker, VH) reactive with β-galactosidase and capable of being expressed in the cytoplasm (Martineau P and Betton J-M: J. Mol. Biol. 292, 921-929 (1999)) was cloned into the aforementioned A7 fiber constructs. This single chain fragment is [SEQ. ID. NO.: 47). To permit cloning into the aforementioned A7 fiber constructs, the single chain fragment was cut with NcoI/EcoR1 from the plasmid pPM163R4 (Martineau P and Betton J-M *supra*) and ligated into a plasmid containing the aforementioned Fiber R7

EGF supplied with a Cla1-Nco1-EcoR1-Xho1 adapter allowing for in-frame ligation of the single chain fragment.

In a continuation of the experiments described in Example 5, expression and cellular localization of recombinant fibers containing the single chain variable chain fragment (VK, linker, VH) reactive with β-galactosidase were evaluated by immunostaining using the following additional primary reagent: β-galactosidase-biotin (SIGMA, G 5025).

In addition to the results shown in Table 2, results results obtained in equivalent experiments for the Anti β-galactosidase single chain Fv fragment fiber construct are shown Table 5:

Table 5: Expression and functional ligand binding in recombinant adenovirus fibers targeted for secretion in COS cells

Fiber	4D2+ aMIg	2A6+ aMIg	β-galactosidase + StrAv-FITC
R7 a-β-galactosidase	+	+	+

In addition to the results shown in Table 3, results obtained in equivalent experiments for the Anti β-galactosidase single chain Fv fragment fiber construct are shown Table 6:

Table 6: Expression and functionality of ligand binding in the nucleus of COS cells in native and selected recombinant fibers after targeting for expression in the cytosol

Fiber	4D2+ aMIg	2A6+ aMIg	β-galactosidase + StrAv-FITC
R7 a-β-galactosidase	+	+	+

EXAMPLE 11:

Phenotypic analysis of fiber proteins.

The following is a further example of modified viruses of the present invention wherein complex polypeptide ligands have been incorporated into a modified adenoviral fibre protein. The example demonstrates the importance for generation of a re-targeted viable and functional Ad vector of two features: (i) the fibre structure modifications should still allow for efficient attachment and cellular entry of the virus, and (ii) ligands inserted into the fibre should be capable of correct folding in the mammalian cell cytoplasm. Solubility is conveniently used to assess correct folding which is typically linked to an absence of disulphide bonds. Preferably the non-native polypeptides described herein will have no more than 2 disulphide bonds, typically no more than 1 and most preferably no disulphide bonds.

Materials and Methods

Cells. HEK-293 cells were obtained from Microbix Inc. (Toronto, Ontario, Canada). Cos7, A431 and Colo205 cells were purchased from the American Type Culture Collection

(ATCC; Manassas, Virginia) and A75 and G43 cells were obtained from Reinder Bolhuis (Daniel Den Hoed Cancer Center, Rotterdam, The Netherlands). All cells were maintained at 37°C and 5% CO_2 in Iscove's medium (Gibco BRL), supplemented with 10% fetal bovine serum (Sigma-Aldrich) and 50 mg/ml Gentamicin (Gibco BRL). Spodoptera frugiperda (Sf9) cells (ATCC) were cultured at 28°C in TC 100 medium (Gibco BRL) with the same supplements as above.

Antibodies. Monoclonal antibody (mAb) against epidermal growth factor (EGF) was purchased from Cambio Ltd. (Cambridge, UK). Antibodies directed against the Vα and Vβ domains of single chain T-cell receptor (scTCR), and anti-idiotypic antibodies directed against mAb G250 (NUH31 and NUH84) were kindly supplied by Reinder Bolhuis. MAb against fiber tail (4D2.5), and mAb against fiber trimer (2A6.36) were obtained from Jeff Engler (University of Alabama at Birmingham, AL). CAR-blocking, fiber knob directed mAb 1D6.14 was supplied by Buck Rogers (UAB at Birmingham, AL). The monoclonal antibody RL2, which is specific for O-linked GlcNAc residues, was obtained from Larry Gerace via Jeff Engler. Biotinylated anti-idiotypic antibodies directed against mAb C242 (Id1, Id13 and Id20) were produced from the original hybridomas. Horse radish peroxidase (HRP)-labeled streptavidin, fluorescein isothiocyanate (FITC)-labeled rabbit anti-mouse immunoglobulin G and streptavidin-FITC were purchased from DAKO (Glostrup, DK).

Generation of recombinant knobless fibers and nomenclature. Recombinant fiber genes were constructed using methods based on ligation, PCR, and splicing by overlap extension (SOE). Gene sequences generated by PCR were sequenced before subcloning. The gene encoding the Ad5 WT fiber was obtained from pAB26 (Microbix, Toronto, Canada) by PCR using the forward primer (SEQ ID NO: 48)

5'-CTC GGA TCC GAT GAA GCG CGC AAG ACC GTC TGA A-3' and reverse primer (SEQ ID NO: 49) 5'-TTC CTC GAG TTA TTC TTG GGC AAT GTA TGA-3' introducing an upstream Bam H I and a downstream Xho I site, respectively. In recombinant fibers, the knob domain was deleted and replaced by a 36 amino acid (aa) extrinsic trimerisation motif derived from the neck region peptide (NRP) of human lung surfactant protein D. The NRP sequence (PDVASLRQQVAELQGQVQHLQAAFSQYKKVELFPNG) (SEQ ID NO:2), followed by a linker sequence from Staphylococcus protein A (Staph-A linker: AKKLNDAQAPKSD), was ligated to the C-terminal end of fiber shaft of different lengths, 1, 7 or 22 repeats.

The resulting constructs were named R1, R7 and R22, respectively. Re-targeting ligands were added to the C-terminal end of the Staph-A linker. For convenient cloning of various ligands, Cla I and Xho I sites were introduced after the linker sequence. All ligands mentioned below were provided with these restriction sites, and the name of the ligand was indicated after the number of shaft repeats in the fiber name. E.g., R1-RGD, R7-EGF, etc (Fig. 4). R7-knob referred to a truncated fiber shaft (repeats 1 to 7), carrying NRP, the Staph-A linker and the natural knob domain, including the last shaft repeat and the shaft-knob junction (Fig. 4). For two R7-knob fiber constructs, an extra linker was inserted into the Nhe I site located on the N-terminal side of NRP. These linkers were derived from the Ad5 WT fiber shaft repeat 17 (L17: TTTACAGCTTCAAACAATTCCAAAAAGCTTGAG) and fiber shaft repeat 22 (L22 GGAAACAAAAATAATGATAAGCTAACTTTGTGTGACC) were named as R7-L17-knob and R7-L22-knob, respectively. The schematic representation of the fibers is shown in Fig 4.

Targeting ligands. Polypeptide ligands. Double stranded

DNA fragments containing the RGD or ACDCRGDCFCG (abbreviated RGD4C) motifs were synthesized as complementary oligonucleotides with single-stranded terminal adapters, and annealed together. The fragment containing the shaft repeat 22, the shaft-knob junction and the entire knob was obtained by PCR from the WT Ad5 fiber gene. The DNA sequence for human EGF was synthesized, cloned and verified by sequencing. The cloned single chain T-cell receptor (which contains disulfide bonds) with specificity towards Mage1/HLA A1 was obtained from Reinder Bolhuis. The scTCR used had the format Vα-Linker-VβCβ.

Monoclonal antibodies. Two mAb single chain fragments (scFv) were used. The first scFv consisted of the variable kappa chain (or VK), a spacing linker, the variable heavy chain (or VH), the joining sequence (JS) and the constant heavy chain domain 2 (or CH2) of mAb G250. MAb G250 has been shown to have a high specificity towards a surface antigen of human renal carcinoma cells.

Affibodies. The second scFv contained the same basic structure VK-linker-JS-VH, and derived from mAb C242, which reacts with colorectal and pancreatic carcinomas. However, to obtain a better selectivity for the designed target cells, a library of C242 scFv-derived affibodies was generated by SOE, using cDNA from the mAb-producing hybridoma cells as the original template. A randomized 18 amino acid peptide ligand $(SNN)_{18}$ was inserted between the VK and VH domains, where S was dA, dG or dC and N was dA, dG, dC or dT. Affibodies were selected by phage display on colorectal carcinoma cells Colo205, and the sequence PPDFVPPAASFPDHSPRG was identified for the peptide ligand. The C242 scFv VK-(PPDFVPPAASFPDHSPRG)-VH was then selected for further work based on its antigen binding ability.

Cellular expression and localization of fiber and ligand reactivity. The genes encoding recombinant fibers were cloned into the vectors pcDNA3.1 and pSecTag2 (Invitrogen BV, Groningen, Germany) for intracellular expression and extracellular release, respectively. Vectors were transfected into Cos7 cells using Lipofectamin (Life Technologies Inc., Gaithersburg, MD, USA) following the manufacturers' instructions. Expression, nuclear transport and functional expression of the fibers with the new cell binding ligands were evaluated by immunostaining as follows. Cells were centrifuged onto microscope slides in a Shandon Cytospin2 cytocentrifuge and air-dried over night at room temperature. The preparations were fixed in 3% paraformaldehyde in PBS and permeabilized with 0.1% Triton-X100 in PBS. After washing in PBS, preparations were incubated with primary antibodies (anti-fiber mAb, or anti-ligand mAb) for 30 min at 37°C in a humid chamber, washed again in PBS and incubated with secondary antibodies labeled with FITC for 30 min at 37°C in a humid chamber. After washing in PBS, preparations were mounted in PBS with 50% glycerol and viewed in a Zeiss Axioskop microscope equipped with appropriate light source and filters for FITC.

Phenotypic analysis of fiber proteins. Recombinant fiber proteins were expressed in insect cells infected with recombinant baculoviruses, and analysed according to four criteria: (i) solubility, (ii) trimerization, (iii) glycosylation and (iv) assembly with recombinant penton base in vivo to form penton capsomers.

(i) For immunological quantification of soluble versus insoluble recombinant fiber fractions, Sf9 cells were lysed in hypotonic buffer (10 mM Tris-HCl buffer, pH 7.5) at 0°C, and the cell lysates were adjusted to isotonic conditions (150 mM NaCl in 10 mM Tris-HCl, pH

7.5) and subjected to centrifugation at 15,000 x g for 10 min. Supernatants and pellets were then analyzed by conventional SDS-PAGE and immunoblotting, using anti-tail 4D2.5 mAb as primary antibody, and [^{35}S]SRL-labeled anti-mouse IgG secondary antibody (Amersham Pharmacia Biotech; 100 μCi/ml; 5 μCi per blot). (ii) Oligomerization status of fiber was assayed by means of non-denaturing SDS-PAGE (referred to as NDS-PAGE) and conventional, denaturing SDS-PAGE. NDS-PAGE differed from SDS-PAGE in that the samples were not denatured by boiling in SDS sample buffer prior to electrophoresis. (iii) Glycosylation of recombinant fibers was assessed both by immunoreaction on blots using the monoclonal antibody RL2 and chemical detection using the DIG Glycan Detection Kit (Roche). (iv) Assembly of fiber with penton base was assayed by co-infecting the same Sf9 cells with two recombinant AcNPV, one expressing the penton base, the other expressing the fiber protein.

The presence of penton capsomer was detected in cell lysates after 40 h post-infection, and analysed by PAGE in native conditions, at low voltage overnight with cooling. Immunological quantification of native penton, penton base and fiber proteins was performed as above, using the corresponding primary antibody (anti-penton base or anti-fiber), followed by [^{35}S]SRL-labeled anti-mouse or anti-rabbit whole IgG secondary antibody. Blots were exposed to radiographic films (Hyperfilm beta-max, Amersham Pharmacia Biotech), and autoradiograms were scanned at 610 nm, using an automatic densitometer (REP-EDC, Helena Laboratories, Beaumont, TX). Alternatively, protein bands were excised from blots and radioactivity measured in a scintillation counter (Beckman LS-6500).

Rescue of recombinant modified fibers into virions.

Recombinant fibers were rescued into the Ad5 genome as described above. Briefly, a 9 kb-fragment from the Ad5 genome, spanning map units (mu) 75.4 to 100 and containing the WT fiber gene, was generated by Spe I and Pac I digestion of pTG3602. The plasmid pTG3602 contained the entire WT Ad5 genome bounded by two Pac I sites. The 75.4-100 mu-fragment was then cloned into pBluescript II SK(-) (Stratagene), after addition of a Pac I site in its MCS, generating the plasmid pGAG9. From pGAG9, a Sac I-Kpn I 3kb-fragment was then subcloned into pBluescript II SK(-). A large deletion downstream of the Nde I site (located within the tail domain of Ad5 fiber) was created in the 3kb-fragment, and the deleted sequence replaced by a Xho I site-containing linker. This generated the plasmid pGAG3, which was the receiving plasmid for all our fiber gene constructs, inserted between Nde I and Xho I sites. pGAG3-inserted recombinant fibers were re-introduced into pGAG9 digested with Nhe I, using homologous recombination in E. coli BJ5183. The resulting recombinant pGAG9 was then excised from the vector with Spe I and Pac I, and joined to the isolated 27 kb-fragment (0-75.5 mu) representing the left-hand segment of the Ad5 genome by Cosmid cloning (SuperCos 1 Cosmid Vector Kit and Gigapack III Gold Packaging Extract, Stratagene). The presence of the correct recombinant fiber in the cosmid clones were verified by PCR and restriction analysis using Spe I and Hind III.

For virus production, recombinant cosmid genomes were isolated after restriction with Pac I, and transfected into cells expressing the receptors corresponding to the fiber-inserted ligand. In standard transfection reactions, 2 μg DNA and 3 μl FuGENE (Roche) were used per 35-mm well, according to the manufacturer's protocol. Ad5 genomes with fiber gene containing either RGD motifs or the WT fiber knob as ligands were

transfected into 293 cells. For the other liganded fibers, Ad5 genomes with the EGF ligand were transfected into A549 cells, scFv-C242 into Colo 205 cells, scFv-G250 into A75 cells and scTCR specific for Mage1/HLA A1 into G43 cells. At least three different transfections were simultaneously performed, each one using a 6-well plate. The occurrence of plaques was determined by microscopic observation of the transfected cell cultures. Verification of recombinant fiber sequence was made by PCR with specific primers for each fiber construct.

Characterization of recombinant fiber and Ad5 virions.

Cellular expression. To evaluate the level of fiber expression, 293 cells were infected with 10 pfu/cell of WT, R7-knob and R7-RGD virus. Cells were harvested and freeze-thawed four times analyzed by SDS-PAGE and western blotting. The blots were reacted with 4D2.5 and revealed with HRP-labeled anti-mouse IgG (DAKO).

Fiber content of Ad5 virions. The fiber copy number of virions was determined, after CsCl purification of Ad5 virions, by SDS-PAGE and western blot analysis as above. The virus loads in acrylamide gels were normalized for equal amounts of infectious particles, determined by virus titration on 293 cells (expressed as plaque forming units/ml; PFU/fml), and for equal amounts of physical particles (PP), determined by protein assay (BioRad). The number of PP was determined by SDS-PAGE analysis and Coomassie blue staining of Ad5 recombinants, co-electrophoresed with a range of standard bovine serum albumin (BSA) samples (2x crystallized, BioRad). Protein content of hexon bands was evaluated by comparison with BSA standard bands, by scanning in an automatic densitometer (REP-EDC, Helena Laboratories, Beaumont, TX). The number of Ad5 PP in

samples was calculated from the mass of 2.91 x 10e-16 g per single virion, i.e. 2.91 mg per 10e+13 virions. The infectivity index represented the ratio of infectious to physical particles.

Growth rate of recombinant Ad5. Growth-rate was measured by the production of plaques in 293 cells. In standard assays, virus in PBS was adsorbed to cell monolayers (4 x 10^4 cells per sample) at 37°C for 1 h. The cells were rinsed once and further incubated in Iscove's medium supplemented with 10 % FCS and 50 mg/ml gentamycin at 37°C. Cells were harvested at 24, 48 and 72 h after infection (pi), centrifuged, dissolved in 0.2 ml PBS and freeze-thawed four times. The supernatants were titered on 293 cells, and the titer expressed as plaque forming units per ml. The total Ad protein content was measured by the IDEIA™ Adenovirus kit (DAKO).

Gene transduction efficiency. Monolayers of 293 cells in 24-well plates were infected as described above with 10 pfu/cell of the recombinant viruses. Cells were harvested at 24 h pi, washed with ice-cold PBS followed by fixation in 0.5 % glutaraldehyde for 15 min. After three washes in PBS, the cells were analysed for transgene GFP expression, using the FL1 emission channel in a FACScan cytometer (Becton-Dickinson, San Jose, CA).

Assay for fiber knob in viral capsids. The presence or absence of accessible knob in virions was assessed by ELISA. Purified virions were diluted with 50 mM carbonate-bicarbonate buffer pH 9.6, to a final concentration of 5 x 10^5 PFU/ml. Aliquots of 100 ml were adsorbed onto ELISA plates overnight at 4°C. Material adsorbed in wells was fixed with 0.5% glutaraldehyde, after which the wells were washed with PBS containing 0.1% Tween-20 (washing buffer), and blocked for 2 h with 200 ml PBS containing 1% BSA. Wells were then washed

three times in washing buffer and incubated for 1 h at room temperature with 100 ml of biotinylated anti-knob mAb 1D6.14 at a concentration of 1 mg/ml. Bound antibody was detected using HRP-streptavidin (DAKO) at a dilution of 1:2,000, for 1 h at room temperature. Colour development was obtained with TMB substrate (CanAg Diagnostics, Göteborg, Sweden), and stopped with 0.12 M HCl for 10 min. Plates were read in a microtiter plate reader set at 450 nm.

Rationale for the construction of recombinant fibers.

Knobless fibers with the extrinsic trimerisation motif (NRP) from human lung surfactant protein D were constructed with different numbers of shaft repeats and different ligand structures. The three constitutive elements of these fiber constructs were considered from a structural and functional point of view: (i) the fiber scaffold, (ii) the flexible linker and (iii) the cell ligands.

Fiber scaffold. The fibers were named according to the number of shaft repeats and the ligand present. As an example, R7-EGF contained the fiber tail, the N-terminal 7 shaft repeats, the NRP motif and the C-terminal EGF peptide as the cellular ligand. R1-RGD fiber had only one shaft repeat (the first one) and RGD as the ligand, etc. Of three possible shaft lengths, short (1 repeat), medium (7 repeats) and long (22 repeats) repeats, the intermediate size fiber with 7 repeats (R7) was chosen as the building scaffold for further constructions and studies with different ligands. For reasons discussed below, the rationale for chosing R7 was based on its high solubility as recombinant protein, the maintenance of most of the fiber biological functions, and on the yields of Ad5/FibR7 virus progeny. The study on the comparative advantages of fiber shafts of different

lengths was performed with R1-RGD, R7-RGD and R22-RGD fibers, three constructs carrying the same cellular ligand RGD.

Linker. To evaluate the possible advantage of an additional linker in the fiber shaft domain, two R7-NRP-knob fibers were constructed with an extra peptide linker inserted between the shaft domain and the NRP motif. Two virus-derived linkers were thus tested, the Ad5 shaft repeat 17 (L17) and the Ad5 shaft repeat 22 (L22). The resulting recombinant fibers were named R7-L17-knob and R7-L22-knob, respectively.

Cellular ligands. Several ligands were tested and compared for functionality. These ligands were designed to re-targeted Ad5 vectors to cell surface molecules of broad distribution, like integrins or HLA molecules, or to less ubiquitous molecules, like malignant cell specific determinants. They varied in size and complexity from a simple tripeptide motif, like RGD, to more elaborated structures, like scTR or scFv, which consist of several polypeptide domains with requirement for proper folding.

Expression of fibers and ligands in mammalian cells. Recombinant fibers were first transiently expressed in mammalian cells, using plasmid-transfected Cos7 cells. The recombinant fibers were expressed from two different vectors, one designed for intracellular expression of recombinant proteins, the other for their extracellular release via the secretory pathway. Cells were assayed for the level of fiber expression, cellular localization and functionality of their ligands. All the different fibers tested exhibited apparent appropriate ligand binding when recovered as secreted proteins. For fibers synthesized from the intracellular expression vector, nuclear localization was observed in all cases, except

for R7-L22-knob. This implied that the cytoplasmic transit and the traverse of the nuclear pore occurred as efficiently for the modified fibers as for WT Ad5 fiber.

However, none of the extrinsic ligands fused to fibers showed the expected binding activity. Even though the intracellularly expressed R7-scTCR fiber could be detected within the nucleus, and was stained with an anti-Va mAb recognizing an epitope independent of the Va domain conformation, no staining could be detected with another, conformation-dependent mAb, directed against the Vb domain. In order to test the possible detrimental effect of the fiber domain on the reactivity of intracellular ligands, our ligands were also expressed as separate constructs from the vector pcDNA. There was no detectable binding of the nonfused intracellular ligands to their specific mAbs, indicating that the ligands behaved similarly when free or fused to the fiber.

There was therefore a profound difference in binding ability between fiber-ligand fusions designed for secretion and designed for cytoplasmic expression. This suggested that fiber-ligand fusion proteins designed to follow the secretory pathway underwent a proper folding and were in the correct conformation. This did not imply that all intracellular fiber-ligand fusion proteins would fold incorrectly, since the cellular environment also seemed to play a major role in this process. A significant difference could be detected between the cytosol of Cos7 and Sf9 cells, in terms of folding pattern and reactivity of fiber-ligand fusion proteins. This was the case for the fiber knob domain of R7-L22-knob, the EGF ligand of R7-EGF and the scFV ligand of R7-G250, which showed some reactivity in Sf9 cells but not in Cos7 cells.

Phenotype of recombinant fibers expressed in insect cells. To further analyze the properties of the liganded fibers, our different constructs were expressed as recombinant proteins in the baculovirus-insect cells system, and assayed for protein solubility, trimerisation, glycosylation and formation of penton by assembly with penton base in vivo in Sf9 cells.

Protein solubility and conformation. The solubility of recombinant proteins is usually considered as a good indicator of their proper folding. We therefore tested all our recombinant fiber proteins for their total expression in insect cells, and determined the proportion recovered in the soluble fraction of the cell lysates. All the fiber recombinants were highly expressed in Sf9 cells, at levels similar to WT Ad5 fiber, although their degree of solubility varied significantly from one to another. We estimated that fibers had a WT-like solubility and thus a proper folding when the soluble fraction contained more than 50 % of the total fiber expressed. A fiber could be considered as misfolded when its soluble fraction represented less than 20 % of the total. According to these criteria, WT, R7-knob, R7-L17-knob, R7-L22-knob, R1-RGD, R7-RGD and R7-RGD4C fibers were mainly recovered in the soluble fraction (60-95 % solubility). In contrast, only 22 % of R22-RGD, and less than 15 % of R7-EGF, R7-C242 and R7-scTCR fiber was found to be soluble, confirming their incorrect folding suggested by the absence of reactivity with their respective mAbs. The least soluble fiber constructs was R7-G250, which was recovered at 95 % in the insoluble fraction (see Fig. 5 and Table 7 below).

Table 7. Phenotypic characterization of recombinant fibers expressed in baculovirus infected Sf9 cells (a)

	Trimerization (b)				
Fiber	Glycosylation	NDS-PAGE	IF	Assembly (c) with penton base	Solubility (d) (%)
WT	+	+	+	+	60
R7-knob	+	+	+	+	72
R7-L17-knob	+	+	+	+	70
R7-L22-knob	+	+	+	+	75
R1-RGD	-	-	+	+	89
R7-RGD	-	+	+	+	94
R22-RGD	-	+	+	+	22
R7-RGD4C	+	+	+	+	78
R7-EGF	-	ND(e)	+	ND	13
R7-C242	-	ND	-	ND	11
R7-G250	-	ND	-	ND	5
R7-scTCR	-	ND	+	ND	12

(a) Baculovirus-infected Sf9 cells were harvested at 48 h after infection and recombinant fibers assayed for different biological functions, solubility, O-GlcNAc glycosylation, trimerisation, and penton capsomer formation.

(b) Trimerisation status of fibers was determined in vitro by electrophoresis of Sf9 cell lysates in SDS-gel without heat denaturation (NDS-PAGE), and in situ, by immunofluorescence staining of fixed cells, using anti-trimer mAb (2A6.36).

(c) The capacity of fibers to assemble with penton base to form penton capsomer was assayed by co-infection of

the same Sf9 cells by two different recombinant baculoviruses, one expressing Ad fiber, the other Ad penton base. Sf9 cell lysates were analysed at 48 h after co-infection for the occurrence of penton (base + fiber), by electrophoresis of native proteins in nondenaturing 6 % acrylamide gels (Karayan et al., 1994)
(d) Solubility was assayed in single baculovirus-infected Sf9 cells. Cells harvested at 48 h after infection were lysed in isotonic buffer, and cell lysates divided in two aliquots. One aliquot was centrifuged at 10,000 x g for 10 min and supernatant was kept, whereas the other was analysed as whole cell lysate. Both aliquots were denatured at 100°C in SDS-sample buffer and electrophoresed in denaturing SDS-gel, and blotted. Blots were reacted with anti-fiber-tail mAb 4D2.5, and radiolabeled secondary antibody, as shown in Fig. 5. Immunoblots were quantitated by autoradiogram scanning and the results were expressed as the percentage of soluble versus total fiber content. Average of three determinations; SD was within 15 % of the reported value for the mean.
(e) ND, not detectable.

Trimerisation. The ability of the different fibers to self-assemble into trimers were electrophoretically determined in vitro by NDS-PAGE analysis of cell lysates, and immunologically in situ by immunofluorescence staining of fixed cells. According to their electrophoretic patterns, R7-EGF, R7-C242, R7-G250 and R7-scTCR were incapable of forming homotrimers, whereas R7-knob, R7-L17-knob R7-L22-knob, R7-RGD, R22-RGD and R7-RGD4C trimerized at WT levels (Table 7). However, immunofluorescence staining of cells using the anti-trimer mAb 2A6.36 showed that all except R7-C242 and R7-G250 formed trimers (Table 7). The apparent discrepancy likely resided in the method of detection, or in the possibility that R7-EGF, R7-C242, R7-G250 and

R7-scTCR fiber trimers became unstable in vitro, or both. Whatever the reason, if fiber trimers formed with a very low efficiency within the cells, immunofluorescence staining was, in this case, more sensitive and more appropriate for trimer detection.

Glycosylation. Likewise, the glycosylation of fibers was analysed by western blotting using RL2, a mAb specific for peptide-linked O-GlcNAc residues. The fiber constructs which reacted with RL2 were, besides the WT, R7-L17-knob, R7-L22-knob, R7-knob and R7-RGD4C (Table 7).

Assembly with penton base. When Sf9 cells were infected with two different baculoviruses, one expressing the Ad fiber, the other the penton base, the two proteins are capable of interacting intracellularly to form penton capsomers. When assayed for this property, the fiber proteins R7-EGF, R7-C242, R7-G250 and R7-scTCR had lost their capacity to assemble with penton base, whereas the other eight recombinant fibers had conserved their assembly function (Table 7).

Rescue of recombinant fiber genes into the viral genome and viability of the viruses. Our recombinant fiber genes were reinserted into the Ad5 viral genome in replacement of the WT fiber gene. Each recombinant viral genome was then introduced by transfection into the corresponding cell line which express the proper receptor for the recombinant fiber. The rescue of viable recombinant viruses was assessed by plaque development. Plaques were observed for the Ad5 genomes harboring the WT fiber, and recombinant fibers R7-knob, R7-L17-knob, R1-RGD, R7-RGD and R7-RGD4C.

Fiber content, infectivity and growth rate of recombinant viruses. The recombinant R7-RGD Ad carrying

seven shaft repeats had been selected as the best choice in terms of fiber trimerisation, stability, assembly with penton base and ligand binding. However, this knobless Ad was less efficient in virus assembly and production, compared to WT virus. In this example, the characteristics of the Ad-R7-knob, which carried the knob domain, were compared to those of WT Ad and knobless Ad-R7-RGD. We first assayed for the production of fibers in 293 cells infected with each of the three Ad at different times pi. The production of fiber proteins from R7-knob Ad was comparable to that of the WT Ad, whereas the level of fiber protein from the R7-RGD Ad was significantly lower.

We then compared the fiber content per virion of the R7-knob Ad with that of the WT Ad. The two viruses exhibited similar amounts of fiber content when normalised to the number of physical particles (PP) but there were more fiber protein in the R7-knob Ad when the two virus samples were normalised to the number of infectious particles (PFU). This indicated that R7-knob Ad was less infectious than the WT. This was confirmed by the infectivity index (PFU : PP ratio), which was calculated from the infectious titer (expressed as PFU/ml, as determined by titration on 293 cells), and from the concentration of PP (determined by biochemical methods). The infectivity index of the Ad-R7-knob (1:100 to 1: 200) was 2- to 8-fold lower than that of WT Ad5 (1:25-1: 50). The infectivity index of Ad-R7-RGD (1 : 200 - 1 : 500) was in a similar range as that of Ad-R7-knob.

The growth rate of the Ad-R7-knob was compared to that of WT virus. Aliquots of cell samples were infected at the same MOI, and infectious virus progeny, determined by plaque assays on 293 cells. The growth curves for the three viruses were similar, but the production of

infectious Ad was inferior to the WT for both Ad-R7-knob and Ad-R7-RGD. Virus uptake was also examined for Ad-R7-knob, and estimated from its ability to transduce the GFP reporter gene into 293 cells, as compared to Ad5-GFP carrying WT fibers (Ad-WTFib). The transducing capacity per PP was found to be 6-fold lower for the R7-knob Ad than for Ad-WTFib.

Receptor-binding capacity of recombinant virions. Ad-FibR7-knob was assayed for its receptor binding capacity, as compared to WT virus. Both WT Ad5 and Ad-FibR7-knob virions were found to react with anti-knob mAb in ELISA, suggesting that the knob epitopes were accessible to mAb on both types of virions. However, the knob immunoreactivity was higher for WT Ad than for Ad-FibR7-knob. This suggested that the cell binding determinants of virions were more efficient or/and had a higher affinity when the knob was carried by a 22-shaft-repeat fiber, as in WT Ad5, than by a shorter, 7-repeat fiber.

These results show that the ligands which were found to be improperly folded within the cytosol and the nucleus all had a conformation that was highly dependent on the formation of disulfide bridges. Thus antibodies or fragments thereof should be used which lack disulfide bridges while retaining native or modified epitope binding.

The number of repeats in the shaft portion should preferably be at least about 7, e.g. 6-12. The presence of a wild type knob (e.g. as one of two types of fibre proteins wherein the second is modified for re-targeting) may be beneficial for generation of recombinant viable Ad virons.

Sequence Listing

[SEQ. ID. NO.: 1] - Adenovirus 5 Fiber Protein encoding sequence.

Sequence Range: 1 to 1746

```
                10         20         30         40         50
atgaagcgcg caagaccgtc tgaagatacc ttcaaccccg tgtatccata
tacttcgcgc gttctggcag acttctatgg aagttggggc acataggtat

                60         70         80         90        100
tgacacggaa accggtcctc caactgtgcc ttttcttact cctccctttg
actgtgcctt tggccaggag gttgacacgg aaaagaatga ggagggaaac

               110        120        130        140        150
tatcccccaa tgggtttcaa gagagtcccc ctggggtact ctctttgcgc
atagggggtt acccaaagtt ctctcagggg gaccccatga gagaaacgcg

               160        170        180        190        200
ctatccgaac ctctagttac ctccaatggc atgcttgcgc tcaaaatggg
gataggcttg gagatcaatg gaggttaccg tacgaacgcg agttttaccc

               210        220        230        240        250
caacggcctc tctctggacg aggccggcaa ccttacctcc caaaatgtaa
gttgccggag agagacctgc tccggccgtt ggaatggagg gttttacatt

               260        270        280        290        300
ccactgtgag cccacctctc aaaaaaacca agtcaaacat aaacctggaa
ggtgacactc gggtggagag ttttttggt tcagtttgta tttggacctt

               310        320        330        340        350
atatctgcac ccctcacagt tacctcagaa gccctaactg tggctgccgc
tatagacgtg gggagtgtca atggagtctt cgggattgac accgacggcg

               360        370        380        390        400
cgcacctcta atggtcgcgg gcaacacact caccatgcaa tcacaggccc
```

gcgtggagat taccagcgcc cgttgtgtga gtggtacgtt agtgtccggg

410 420 430 440 450

cgctaaccgt gcacgactcc aaacttagca ttgccaccca aggacccctc

gcgattggca cgtgctgagg tttgaatcgt aacggtgggt tcctggggag

460 470 480 490 500

acagtgtcag aaggaaagct agccctgcaa acatcaggcc ccctcaccac

tgtcacagtc ttcctttcga tcgggacgtt tgtagtccgg gggagtggtg

510 520 530 540 550

caccgatagc agtaccctta ctatcactgc ctcacccccc ctaactactg

gtggctatcg tcatgggaat gatagtgacg gagtgggggg gattgatgac

560 570 580 590 600

ccactggtag cttgggcatt gacttgaaag agcccattta tacacaaaat

ggtgaccatc gaacccgtaa ctgaactttc tcgggtaaat atgtgtttta

610 620 630 640 650

ggaaaactag gactaaagta cggggctcct ttgcatgtaa cagacgacct

ccttttgatc ctgatttcat gccccgagga aacgtacatt gtctgctgga

660 670 680 690 700

aaacactttg accgtagcaa ctggtccagg tgtgactatt aataatactt

tttgtgaaac tggcatcgtt gaccaggtcc acactgataa ttattatgaa

710 720 730 740 750

ccttgcaaac taaagttact ggagccttgg gttttgattc acaaggcaat

ggaacgtttg atttcaatga cctcggaacc caaaactaag tgttccgtta

760 770 780 790 800

atgcaactta atgtagcagg aggactaagg attgattctc aaaacagacg

tacgttgaat tacatcgtcc tcctgattcc taactaagag ttttgtctgc

810 820 830 840 850

ccttatactt gatgttagtt atccgtttga tgctcaaaac caactaaatc

ggaatatgaa ctacaatcaa taggcaaact acgagttttg gttgatttag

860 870 880 890 900
taagactagg acagggccct ctttttataa actcagccca caacttggat
attctgatcc tgtcccggga gaaaaatatt tgagtcgggt gttgaaccta

910 920 930 940 950
attaactaca acaaaggcct ttacttgttt acagcttcaa acaattccaa
taattgatgt tgtttccgga aatgaacaaa tgtcgaagtt tgttaaggtt

960 970 980 990 1000
aaagcttgag gttaacctaa gcactgccaa ggggttgatg tttgacgcta
tttcgaactc caattggatt cgtgacggtt ccccaactac aaactgcgat

1010 1020 1030 1040 1050
cagccatagc cattaatgca ggagatgggc ttgaatttgg ttcacctaat
gtcggtatcg gtaattacgt cctctacccg aacttaaacc aagtggatta

1060 1070 1080 1090 1100
gcaccaaaca caaatcccct caaaacaaaa attggccatg gcctagaatt
cgtggtttgt gtttagggga gttttgtttt taaccggtac cggatcttaa

1110 1120 1130 1140 1150
tgattcaaac aaggctatgg ttcctaaact aggaactggc cttagttttg
actaagtttg ttccgatacc aaggatttga tccttgaccg gaatcaaaac

1160 1170 1180 1190 1200
acagcacagg tgccattaca gtaggaaaca aaaataatga taagctaact
tgtcgtgtcc acggtaatgt catcctttgt ttttattact attcgattga

1210 1220 1230 1240 1250
ttgtggacca caccagctcc atctcctaac tgtagactaa atgcagagaa
aacacctggt gtggtcgagg tagaggattg acatctgatt tacgtctctt

1260 1270 1280 1290 1300
agatgctaaa ctcactttgg tcttaacaaa atgtggcagt caaatacttg
tctacgattt gagtgaaacc agaattgttt tacaccgtca gtttatgaac

1310 1320 1330 1340 1350
ctacagtttc agttttggct gttaaaggca gtttggctcc aatatctgga

gatgtcaaag tcaaaaccga caatttccgt caaaccgagg ttatagacct

1360 1370 1380 1390 1400

acagttcaaa gtgctcatct tattataaga tttgacgaaa atggagtgct

tgtcaagttt cacgagtaga ataatattct aaactgcttt tacctcacga

1410 1420 1430 1440 1450

actaaacaat tccttcctgg acccagaata ttggaacttt agaaatggag

tgatttgtta aggaaggacc tgggtcttat aaccttgaaa tctttacctc

1460 1470 1480 1490 1500

atcttactga aggcacagcc tatacaaacg gtgttggatt tatgcctaac

tagaatgact tccgtgtcgg atatgtttgc cacaacctaa atacggattg

1510 1520 1530 1540 1550

ctatcagctt atccaaaatc tcacggtaaa actgccaaaa gtaacattgt

gatagtcgaa taggttttag agtgccattt tgacggtttt cattgtaaca

1560 1570 1580 1590 1600

cagtcaagtt tacttaaacg gagacaaaac taaacctgta acactaacca

gtcagttcaa atgaatttgc ctctgttttg atttggacat tgtgattggt

1610 1620 1630 1640 1650

ttacactaaa cggtacacag gaaacaggag acacaactcc aagtgcatac

aatgtgattt gccatgtgtc ctttgtcctc tgtgttgagg ttcacgtatg

1660 1670 1680 1690 1700

tctatgtcat tttcatggga ctggtctggc cacaactaca ttaatgaaat

agatacagta aaagtaccct gaccagaccg gtgttgatgt aattacttta

1710 1720 1730 1740

atttgccaca tcctcttaca ctttttcata cattgcccaa gaataa

taaacggtgt aggagaatgt gaaaaagtat gtaacgggtt cttatt

[SEQ. ID. NO.: 2] - Neck Region Peptide from Human Lung Surfactant Protein D.

PDVASLRQQVEALQGQVQHLQAAFSQYKKVELFPNG

[SEQ. ID. NO.: 3] - Recombinant Fiber A1

gaattcg atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat
cca tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc
ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct ttg
cgc cta tcc gaa cct cta gtt acc tcc aat ggc atg cct gac gta gca agc
tta cga caa cag gta gaa gcc ttg caa ggg cag gta caa cac tta cag gcg
gca ttt agc caa tac aaa aag gta gag ttg ttt cca aac gga gcc aag aag
ctg aac gac gcc cag gcc ccc aag agc gac cca tcg atc taa ctc gag

[SEQ. ID. NO.: 4] - Human epidermal growth factor (EGF)

Sequence Range: 1 to 183

```
                10         20         30         40         50
        gggaattcat gaactccgac tccgaatgtc cattgtccca cgacggttac
        cccttaagta cttgaggctg aggcttacag gtaacagggt gctgccaatg
                M   N  S  D   S  E  C   P  L  S  H   D  G  Y>

                60         70         80         90        100
        tgtttgcacg acggtgtttg tatgtacatc gaagctttgg acaagtacgc
        acaaacgtgc tgccacaaac atacatgtag cttcgaaacc tgttcatgcg
         C  L  H   D  G  V  C   M  Y  I   E  A  L   D  K  Y  A>

               110        120        130        140        150
        ttgtaactgt gttgttggtt acatcggtga aagatgtcaa tacagagact
        aacattgaca caacaaccaa tgtagccact ttctacagtt atgtctctga
          C  N  C   V  V  G   Y  I  G  E   R  C  Q   Y  R  D>
```

```
        160         170         180
tgaagtggtg ggaattgaga tgataagaat tcc
acttcaccac ccttaactct actattctta agg
L  K  W  W   E  L  R   *>
```

[SEQ. ID. NO.: 5] - Amino acid linker from Staphylococcal protein A

Sequence Range: 1 to 39

```
          10          20          30
gccaagaagc tgaacgacgc ccaggccccc aagagcgac
cggttcttcg acttgctgcg ggtccggggg ttctcgctg
 A  K  K   L  N  D  A   Q  A  P   K  S  D>
_________a____SPA 48-60_____a_________
```

[SEQ. ID. NO.: 6] - Fiber construct A1 EGF

Sequence Range: 1 to 513

```
          10          20          30          40          50
atgaagcgcg caagaccgtc tgaagatacc ttcaaccccg tgtatccata
tacttcgcgc gttctggcag acttctatgg aagttggggc acataggtat
 M  K  R   A  R  P  S   E  D  T   F  N  P   V  Y  P  Y>
_________b_________b__TAIL___b_________b_________>

          60          70          80          90          100
tgacacggaa accggtcctc caactgtgcc ttttcttact cctccctttg
actgtgcctt tggccaggag gttgacacgg aaaagaatga ggagggaaac
  D  T  E   T  G  P   P  T  V  P   F  L  T   P  P  F>
_________b_________b__TAIL___b_________b_________>

         110         120         130         140         150
tatcccccaa tgggtttcaa gagagtcccc ctggggtact ctctttgcgc
atagggggtt acccaaagtt ctctcagggg gaccccatga gagaaacgcg
                                              L   S  L  R>
```

```
                                                        __a_________>
V  S  P  N   G  F  Q   E  S  P   P  G  V>
_________b_____TAIL_________b_______>

      160        170        180        190        200
ctatccgaac ctctagttac ctccaatggc atgcctgacg tagcaagctt
gataggcttg gagatcaatg gaggttaccg tacggactgc atcgttcgaa
 L  S  E   P  L  V  T   S  N  G   M>
_________a____R1__a_________a__>
                                          P  D   V  A  S  L>
                                       ___LUNG SURFAC____>

      210        220        230        240        250
acgacaacag gtagaagcct tgcaagggca ggtacaacac ttacaggcgg
tgctgttgtc catcttcgga acgttcccgt ccatgttgtg aatgtccgcc
  R  Q  Q   V  E  A   L  Q  G  Q   V  Q  H   L  Q  A>
__________LUNG SURFACTANT PROTEIN D TRIM__c_________>

      260        270        280        290        300
catttagcca atacaaaaag gtagagttgt ttccaaacgg agccaagaag
gtaaatcggt tatgtttttc catctcaaca aaggtttgcc tcggttcttc
A  F  S  Q   Y  K  K   V  E  L   F  P  N  G>
_______LUNG SURFACTANT PROTEIN D TRIM_____c_>
                                             A  K  K>
                                          _________>

                                    >ClaI
                                      |
      310        320        330      |  340        350
ctgaacgacg cccaggcccc caagagcgac ccatcgatca tgaactccga
gacttgctgc gggtccgggg gttctcgctg ggtagctagt acttgaggct
                                             M  N  S  D>
                                          _d_________>
 L  N  D   A  Q  A  P   K  S  D>
__________SPA 48-60_e_________>
                                       P  S  I>
                                    _________>
```

```
          360        370        380        390        400
        ctccgaatgt ccattgtccc acgacggtta ctgtttgcac gacggtgttt
        gaggcttaca ggtaacaggg tgctgccaat gacaaacgtg ctgccacaaa
          S  E  C   P  L  S   H  D  G  Y   C  L  H   D  G  V>
        ___EPIDERMAL GROWTH FACTOR (UROGASTRONE); CODON_ST____>

          410        420        430        440        450
        gtatgtacat cgaagctttg gacaagtacg cttgtaactg tgttgttggt
        catacatgta gcttcgaaac ctgttcatgc gaacattgac acaacaacca
        C  M  Y  I   E  A  L   D  K  Y   A  C  N  C   V  V  G>
        ___EPIDERMAL GROWTH FACTOR (UROGASTRONE); CODON_ST____>

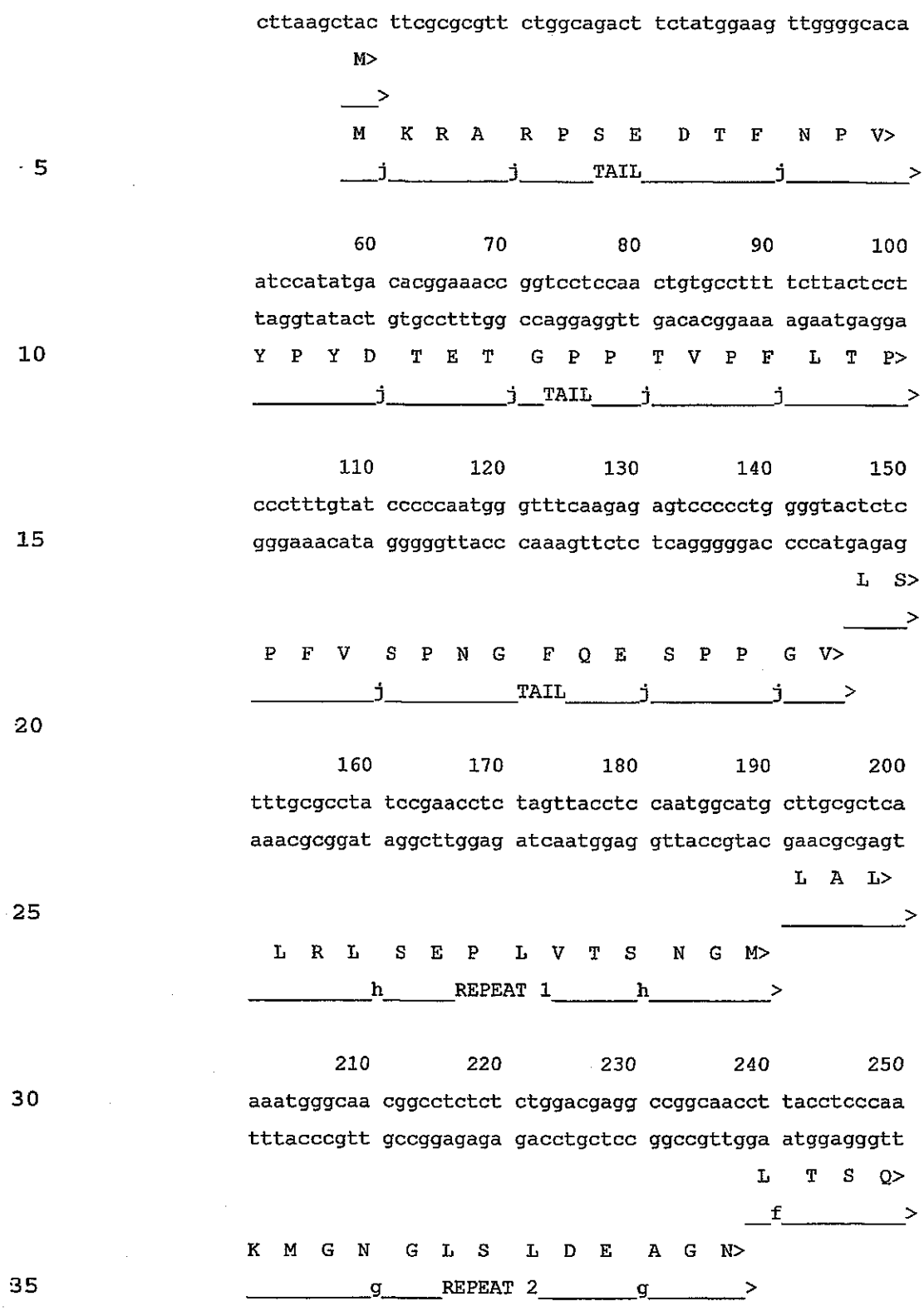
          460        470        480        490        500
        tacatcggtg aaagatgtca atacagagac ttgaagtggt gggaattgag
        atgtagccac tttctacagt tatgtctctg aacttcacca cccttaactc
         Y  I  G   E  R  C  Q   Y  R  D   L  K  W   W  E  L  R>
        ___EPIDERMAL GROWTH FACTOR (UROGASTRONE); CODON_ST____>

          >xho1
            |
            |  510
        atgactcgag ggg
        tactgagctc ccc
          X>
        ___>
```

[SEQ. ID. NO.: 7] - Fiber Construct A7 EGF

Sequence Range: 1 to 811

```
        >EcoR1
          |
          |      10         20         30         40         50
          gaattcgatg aagcgcgcaa gaccgtctga agataccttc aacccctgt
```


cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
M>
___>
M K R A R P S E D T F N P V>
___j_____j_____TAIL_____j_____>
60 70 80 90 100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y P Y D T E T G P P T V P F L T P>
_____j_____j__TAIL___j_____j_____>
110 120 130 140 150
ccctttgtat cccccaatgg gtttcaagag agtccccctg gggtactctc
gggaaacata gggggttacc caaagttctc tcagggggac cccatgagag
L S>
_____>
P F V S P N G F Q E S P P G V>
_____j_____TAIL_____j_____j_____>
160 170 180 190 200
tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca
aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt
L A L>
_____>
L R L S E P L V T S N G M>
_____h_____REPEAT 1_____h_____>
210 220 230 240 250
aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa
tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt
L T S Q>
__f_____>
K M G N G L S L D E A G N>
_____g_____REPEAT 2_____g_____>
5
10
15
20
25
30
35

```
          260        270        280        290        300
aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa
ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt
                                                  I  N>
                                                  _____>
 N  V  T   T  V  S  P   P  L  K   K  T  K   S  N>
_________f________REPEAT 3 ____f_________f_____>

          310        320        330        340        350
cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg
ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc
                                                L  T  V>
                                                _________>
  L  E  I   S  A  P   L  T  V  T   S  E  A>
_________e_____REPEAT 4______e_________>

          360        370        380        390        400
ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                          L  T   M  Q  S>
                                          ___REPEAT 6_____>
A  A  A  A   P  L  M   V  A  G   N  T>
_________d___REPEAT 5_________d____>

          410        420        430        440        450
caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                    L  S  I   A  T  Q  G>
                                    _____REPEAT 7______>
 Q  A  P   L  T  V  H   D  S  K>
__________REPEAT 6__c_________>

          460        470        480        490        500
acccctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
```

```
                                                    L  A>
                                                   _____a_>
                                  P  L  T   V  S  E   G  K>
                                 _________REPEAT 7____b_____>
                                                                    P  D  V   A  S  L>
                                                                   ___LUNG SURFACTA____>
                                       510         520         530         540         550
                                 gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
                                 ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
                                 R  Q  Q  V   E  A  L   Q  G  Q   V  Q  H  L   Q  A  A>
                                 __________LUNG SURFACTANT PROTEIN D TRIM__k__________>

                                       560         570         580         590         600
                                 tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
                                 aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
                                  F  S  Q   Y  K  K  V   E  L  F   P  N  G>
                                 _____LUNG SURFACTANT PROTEIN D TRIM_______>
                                                                          A  K  K  L>
                                                                        _m_________>

                                                          >ClaI
                                                            |
                                       610         620         630 |       640         650
                                 gaacgacgcc caggccccca agagcgaccc atcgatcatg aactccgact
                                 cttgctgcgg gtccgggggt tctcgctggg tagctagtac ttgaggctga
                                                                       M   N  S  D>
                                                                     ___l_________>
                                   N  D  A   Q  A  P   K  S  D>
                                 __________SPA 48-60__m_______>
                                                                 P   S  I>
                                                               __n_______>

                                       660         670         680         690         700
                                 ccgaatgtcc attgtcccac gacggttact gtttgcacga cggtgtttgt
                                 ggcttacagg taacagggtg ctgccaatga caaacgtgct gccacaaaca
```

```
                S  E  C  P   L  S  H   D  G  Y   C  L  H  D   G  V  C>
               ___EPIDERMAL GROWTH FACTOR (UROGASTRONE); CODON_ST____>

                      710        720        730        740        750
5              atgtacatcg aagctttgga caagtacgct tgtaactgtg ttgttggtta
               tacatgtagc ttcgaaacct gttcatgcga acattgacac aacaaccaat
                M  Y  I   E  A  L  D   K  Y  A   C  N  C   V  V  G  Y>
               ___EPIDERMAL GROWTH FACTOR (UROGASTRONE); CODON_ST____>
                      760        770        780        790        800
10             catcggtgaa agatgtcaat acagagactt gaagtggtgg gaattgagat
               gtagccactt tctacagtta tgtctctgaa cttcaccacc cttaactcta
                 I  G  E   R  C  Q   Y  R  D  L   K  W  W   E  L  R>
               ___EPIDERMAL GROWTH FACTOR (UROGASTRONE); CODON_ST____>

15             >xhol
                 |
                 |   810
               gactcgaggg g
               ctgagctccc c
20             X>
               _>
```

[SEQ. ID. NO.: 8] - G250 Construct

25 Sequence Range: 1 to 1056

```
                      10         20         30         40         50
               gacattgtga tgacccagtc tcaaagattc atgtccacaa cagtaggaga
               ctgtaacact actgggtcag agtttctaag tacaggtgtt gtcatcctct
30              D  I  V   M  T  Q  S   Q  R  F   M  S  T   T  V  G  D>
               _________c_________c___VK____c_________c_________>

                      60         70         80         90         100
               cagggtcagc atcacctgca aggccagtca gaatgtggtt tctgctgttg
35             gtcccagtcg tagtggacgt tccggtcagt cttacaccaa agacgacaac
```

```
  R  V  S  I  T  C  K  A  S  Q  N  V  V  S  A  V>
_____c_____c___VK____c_____c_____>

      110        120        130        140        150
cctggtatca acagaaacca ggacaatctc ctaaactact gatttactca
ggaccatagt tgtctttggt cctgttagag gatttgatga ctaaatgagt
A  W  Y  Q  Q  K  P  G  Q  S  P  K  L  L  I  Y  S>
_____c_____c___VK____c_____c_____>

      160        170        180        190        200
gcatccaatc ggtacactgg agtccctgat cgcttcacag gcagtggatc
cgtaggttag ccatgtgacc tcagggacta gcgaagtgtc cgtcacctag
 A  S  N  R  Y  T  G  V  P  D  R  F  T  G  S  G  S>
_____c_____c___VK____c_____c_____>

      210        220        230        240        250
tgggacagat ttcactctca ccattagcaa tatgcagtct gaagacctgg
accctgtcta aagtgagagt ggtaatcgtt atacgtcaga cttctggacc
  G  T  D  F  T  L  T  I  S  N  M  Q  S  E  D  L>
_____c_____c___VK____c_____c_____>

      260        270        280        290        300
ctgatttttt ctgtcaacaa tatagcaact atccgtggac gttcggtgga
gactaaaaaa gacagttgtt atatcgttga taggcacctg caagccacct
A  D  F  F  C  Q  Q  Y  S  N  Y  P  W  T  F  G  G>
_____c_____c___VK____c_____c_____>

      310        320        330        340        350
ggcaccaagc tggaaatcaa aggatctggc tctacttccg gtagcggcaa
ccgtggttcg acctttagtt tcctagaccg agatgaaggc catcgccgtt
 G  T  K  L  E  I  K>
_____VK_____c_>
                     G  S  G  S  T  S  G  S  G  K>
                  _____d___LINK___d_____>
```

```
       360        370        380        390        400
atcctctgaa ggcaaaggta ctagagacgt gaagctcgtg gagtctgggg
taggagactt ccgtttccat gatctctgca cttcgagcac ctcagacccc
                                D  V   K  L  V   E  S  G>
                                _____a_____VH__a_________>
  S  S  E   G  K  G   T  R>
__________LINK_____d_____>

       410        420        430        440        450
gaggcttagt gaagcttgga gggtccctga aactctcctg tgcagcctct
ctccgaatca cttcgaacct cccagggact ttgagaggac acgtcggaga
G  G  L  V   K  L  G   G  S  L   K  L  S  C   A  A  S>
__________a_________a___VH____a_________a_________>

       460        470        480        490        500
ggattcactt tcagtaacta ttacatgtct tgggttcgcc agactccaga
cctaagtgaa agtcattgat aatgtacaga acccaagcgg tctgaggtct
 G  F  T   F  S  N  Y   Y  M  S   W  V  R   Q  T  P  E>
__________a_________a___VH____a_________a_________>

       510        520        530        540        550
gaagaggctg gagttggtcg cagccattaa tagtgatggt ggtatcacct
cttctccgac ctcaaccagc gtcggtaatt atcactacca ccatagtgga
  K  R  L   E  L  V   A  A  I  N   S  D  G   G  I  T>
__________a_________a___VH____a_________a_________>

       560        570        580        590        600
actatctaga cactgtgaag ggccgattca ccatttcaag agacaatgcc
tgatagatct gtgacacttc ccggctaagt ggtaaagttc tctgttacgg
Y  Y  L  D   T  V  K   G  R  F   T  I  S  R   D  N  A>
__________a_________a___VH____a_________a_________>

       610        620        630        640        650
aagaacaccc tgtacctgca aatgagcagt ctgaagtctg aggacacagc
ttcttgtggg acatggacgt ttactcgtca gacttcagac tcctgtgtcg
```

K N T L Y L Q M S S L K S E D T A>
_____a_____a___VH____a_____a_____>

660 670 680 690 700
cttgttttac tgtgcaagac accgctcggg ctacttttct atggactact
gaacaaaatg acacgttctg tggcgagccc gatgaaaaga tacctgatga
L F Y C A R H R S G Y F S M D Y>
_____a_____a___VH____a_____a_____>

710 720 730 740 750
ggggtcaagg aacctcagtc accgtctcct catgcccacc gtgcccagca
ccccagttcc ttggagtcag tggcagagga gtacgggtgg cacgggtcgt
W G Q G T S V T V S S>
_____a____VH___a_____a__>

A>
___>
C P P C P>
____HINGE_____>

760 770 780 790 800
cctgaactcc taggggacc gtcagtcttc ctcttccccc caaaacccaa
ggacttgagg atccccctgg cagtcagaag gagaaggggg gttttgggtt
P E L L G G P S V F L F P P K P K>
_____b_____b___CH2___b_____b_____>

810 820 830 840 850
ggacaccctc atgatctccc ggacccctga ggtcacatgc gtggtggtgg
cctgtgggag tactagaggg cctggggact ccagtgtacg caccaccacc
D T L M I S R T P E V T C V V V>
_____b_____b___CH2___b_____b_____>

860 870 880 890 900
acgtgagcca cgaagaccct gaggtcaagt tcaactggta cgtggacggc
tgcactcggt gcttctggga ctccagttca agttgaccat gcacctgccg
D V S H E D P E V K F N W Y V D G>

```
          _____b_____b__CH2__b_____b_____>

                910        920        930        940        950
          gtggaggtgc ataatgccaa gacaaagccg cgggaggagc agtacaacag
          cacctccacg tattacggtt ctgtttcggc gccctcctcg tcatgttgtc
           V  E  V   H  N  A  K   T  K  P   R  E  E   Q  Y  N  S>
          _____b_____b__CH2__b_____b_____>

                960        970        980        990       1000
          cacgtaccgg gtggtcagcg tcctcaccgt cctgcaccag gactggctga
          gtgcatggcc caccagtcgc aggagtggca ggacgtggtc ctgaccgact
            T  Y  R   V  V  S   V  L  T  V   L  H  Q   D  W  L>
          _____b_____b__CH2__b_____b_____>

               1010       1020       1030       1040       1050
          atggcaagga gtacaagtgc aaggtctcca acaaagccct cccagccccc
          taccgttcct catgttcacg ttccagaggt tgtttcggga gggtcggggg
          N  G  K  E   Y  K  C   K  V  S   N  K  A  L   P  A  P>
          _____b_____b__CH2__b_____b_____>

          atcgag
          tagctc
           I  E>
          _____>
```

[SEQ. ID. NO.: 9] - C242 Construct

Sequence Range: 1 to 793

```
          >Sfi        >EcoR1   >Cla1
           |            |       |
           |    10      |20     |   30        40        50
          gcggcccagc cggccacgaa ttcatcgatg gatattgtga tgactcaggc
```

```
                cgccgggtcg gccggtgctt aagtagctac ctataacact actgagtccg
                   A  Q   P  A  T  N   S  S  M>
                  _______b_BACKPRIMER_________>
                                                D  I  V   M  T  Q  A>
5                                               ____VK PHARMACIA_____>
                A>
                __>

                       60         70         80         90         100
10              tgcaccctct gtacctgtca ctcctggaga gtcagtatcc atctcctgca
                acgtgggaga catggacagt gaggacctct cagtcatagg tagaggacgt
                  A  P  S   V  P  V   T  P  G  E   S  V  S   I  S  C>
                _________d________VK PHARMACIA_________d________>

15                    110        120        130        140        150
                ggtctagtaa gagtctcctg catagtaatg gcaacactta cttgtattgg
                ccagatcatt ctcagaggac gtatcattac cgttgtgaat gaacataacc
                R  S  S  K   S  L  L   H  S  N   G  N  T  Y   L  Y  W>
                _________d________VK PHARMACIA_________d________>
20                       S  L  L   H  S  N   G  N  T  Y   L  Y>
                         ________f______CDR1________f______>

                      160        170        180        190        200
                ttcctgcaga ggccaggcca gtctcctcag ctcctgatat atcggatgtc
25              aaggacgtct ccggtccggt cagaggagtc gaggactata tagcctacag
                 F  L  Q   R  P  G  Q   S  P  Q   L  L  I   Y  R  M  S>
                _________d________VK PHARMACIA_________d________>
                                                              R  M  S>
                                                             _______>
30
                      210        220        230        240        250
                caaccttgtc tcaggagtcc cagacaggtt cagtggcagt gggtcaggaa
                gttggaacag agtcctcagg gtctgtccaa gtcaccgtca cccagtcctt
                  N  L  V   S  G  V   P  D  R  F   S  G  S   G  S  G>
35              _________d________VK PHARMACIA_________d________>
```

```
  N  L  V   S>
_________e___>

       260        270        280        290        300
ctgctttcac actgagaatc agtagagtgg aggctgagga tgtgggtgtt
gacgaaagtg tgactcttag tcatctcacc tccgactcct acacccacaa
T  A  F  T   L  R  I   S  R  V   E  A  E  D   V  G  V>
_________d________VK PHARMACIA_________d_________>

       310        320        330        340        350
tattactgtc tgcaacatct agagtatccg ttcacgttcg gtcctgggac
ataatgacag acgttgtaga tctcataggc aagtgcaagc caggaccctg
             L  Q  H  L   E  Y  P   F  T>
          _a_________CDR3_______a_____>
 Y  Y  C   L  Q  H  L   E  Y  P   F  T  F   G  P  G  T>
_________d________VK PHARMACIA_________d_________>

       360        370        380        390        400
caagctggag ctgaaacggc ccccggactt tgttcccccg gccgctagtt
gttcgacctc gactttgccg ggggcctgaa acaaggggggc cggcgatcaa
  K  L  E   L  K  R>
___VK PHARMACIA_____>
                        P  P  D  F   V  P  P   A  A  S>
                     _l_________l__825_____l_________>

       410        420        430        440        450
tccctgatca ctcccctcgt ggccaggtcc agttggtgca gtctggacct
agggactagt gaggggagca ccggtccagg tcaaccacgt cagacctgga
                              Q  V   Q  L  V   Q  S  G  P>
                           ____VH PHARMACIA [SPLIT]_____>
F  P  D  H   S  P  R   G>
_________825_______l___>

       460        470        480        490        500
gagctgaaga agcctggaga gacagtcaag atctcctgca aggcttctga
```

```
ctcgacttct tcggacctct ctgtcagttc tagaggacgt tccgaagact
 E  L  K   K  P  G  E   T  V  K   I  S  C   K  A  S  D>
__________h_____VH PHARMACIA [SPLIT]_______h__________>

      510        520        530        540        550
ttataccttc acatactatg gaatgaactg ggtgaagcag gctccgggaa
aatatggaag tgtatgatac cttacttgac ccacttcgtc cgaggccctt
  Y  T  F   T  Y  Y   G  M  N  W   V  K  Q   A  P  G>
__________h_____VH PHARMACIA [SPLIT]_______h__________>
                 Y  Y   G  M  N>
                _____CDR1_______>

      560        570        580        590        600
agggtttaaa gtggatgggc tggatagaca ccaccactgg agagccaaca
tcccaaattt cacctacccg acctatctgt ggtggtgacc tctcggttgt
K  G  L  K   W  M  G   W  I  D   T  T  T  G   E  P  T>
__________h_____VH PHARMACIA [SPLIT]_______h__________>
                       W  I  D   T  T  T  G   E  P  T>
                      _________CDR2 [SPLIT]___________>

      610        620        630        640        650
tatgctgaag attttaaggg acggattgcc ttctctttgg agacctctgc
atacgacttc taaaattccc tgcctaacgg aagagaaacc tctggagacg
 Y  A  E   D  F  K  G   R  I  A   F  S  L   E  T  S  A>
__________h_____VH PHARMACIA [SPLIT]_______h__________>
 Y  A  E   D  F  K  G>
_____CDR2 [SPLIT]____i_>

      660        670        680        690        700
cagcactgcc tatttgcaga tcaaaaacct caaaaatgag gacacggcta
gtcgtgacgg ataaacgtct agtttttgga gttttttactc ctgtgccgat
  S  T  A   Y  L  Q   I  K  N  L   K  N  E   D  T  A>
__________h_____VH PHARMACIA [SPLIT]_______h__________>

      710        720        730        740        750
```

```
                catatttctg tgcaagacgg gggccttaca actggtactt tgatgtctgg
                gtataaagac acgttctgcc cccggaatgt tgaccatgaa actacagacc
                                  R   G   P   Y   N   W   Y   F   D  V>
                                 ___g_________CDR3_______g_______>
5               T   Y   F   C   A   R   R   G   P   Y   N   W   Y   F   D   V   W>
                _________h_____VH PHARMACIA [SPLIT]_______h_________>

                                      >Xho(-)StopXho_adapt_[Split]
                                                                |
10                    760         770         780             |       790
                ggccaaggga ccacggtcac cgtctcctca ctcgattaac tcg
                ccggttccct ggtgccagtg gcagaggagt gagctaattg agc
                                                  L   X>
                                                 _____>
15               G   Q   G   T   T   V   T   V   S   S>
                _____VH PHARMACIA [SPLIT]_______>
```

[SEQ. ID. NO.: 10] - Fiber Construct A1 G250

Sequence Range: 1 to 1411

```
                >EcoR1
                  |
                  |       10         20         30         40         50
                  gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
                  cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
                         M>
                        ___>
                         M   K   R   A   R   P   S   E   D   T   F   N   P   V>
                        ___f_________f_____TAIL_________f_________>

                         60         70         80         90         100
                  atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
                  taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
                  Y   P   Y   D   T   E   T   G   P   P   T   V   P   F   L   T   P>
                  _________f_________f__TAIL___f_________f_________>

                        110        120        130        140        150
                  ccctttgtat cccccaatgg gtttcaagag agtccccctg gggtactctc
                  gggaaacata gggggttacc caaagttctc tcaggggggac cccatgagag
                                                                  L   S>
                                                                 _____>
                   P   F   V   S   P   N   G   F   Q   E   S   P   P   G   V>
                  _________f_________TAIL_____f_________f_____>

                        160        170        180        190        200
                  tttgcgccta tccgaacctc tagttacctc caatggcatg cctgacgtag
                  aaacgcggat aggcttggag atcaatggag gttaccgtac ggactgcatc
                    L   R   L   S   E   P   L   V   T   S   N   G   M>
                  _________d_____REPEAT 1______d_________>
                                                                  P   D   V>
```

```
                                                          _________>

          210         220         230         240         250
caagcttacg acaacaggta gaagccttgc aagggcaggt acaacactta
gttcgaatgc tgttgtccat cttcggaacg ttcccgtcca tgttgtgaat
A  S  L  R   Q  Q  V   E  A  L   Q  G  Q  V   Q  H  L>
__________LUNG SURFACTANT PROTEIN D TRIM__g_________>

          260         270         280         290         300
caggcggcat ttagccaata caaaaaggta gagttgtttc caaacggagc
gtccgccgta aatcggttat gtttttccat ctcaacaaag gtttgcctcg
                                                         A>
                                                         __>
 Q  A  A   F  S  Q  Y   K  K  V   E  L  F   P  N  G>
__________LUNG SURFACTANT PROTEIN D TRIM___g_______>

                                                  >Cla1
                                                    |
          310         320         330         340         350
caagaagctg aacgacgccc aggcccccaa gagcgaccca tcgatcgaca
gttcttcgac ttgctgcggg tccgggggtt ctcgctgggt agctagctgt
  K  K  L   N  D  A   Q  A  P  K   S  D>
__________b____SPA 48-60_______b_______>
                                                   P   S  I>
                                                  ___c_____>
                                                          D>
                                                          ____>

          360         370         380         390         400
ttgtgatgac ccagtctcaa agattcatgt ccacaacagt aggagacagg
aacactactg ggtcagagtt tctaagtaca ggtgttgtca tcctctgtcc
I  V  M  T   Q  S  Q   R  F  M   S  T  T  V   G  D  R>
__________j_________j___VK____j_________j_________>

          410         420         430         440         450
```

```
gtcagcatca cctgcaaggc cagtcagaat gtggtttctg ctgttgcctg
cagtcgtagt ggacgttccg gtcagtctta caccaaagac gacaacggac
 V  S  I   T  C  K  A   S  Q  N   V  V  S   A  V  A  W>
_________j_________j___VK____j_________j_________>
        460         470         480         490         500
gtatcaacag aaaccaggac aatctcctaa actactgatt tactcagcat
catagttgtc tttggtcctg ttagaggatt tgatgactaa atgagtcgta
  Y  Q  Q   K  P  G   Q  S  P   K  L  L   I  Y  S  A>
_________j_________j___VK____j_________j_________>

        510         520         530         540         550
ccaatcggta cactggagtc cctgatcgct tcacaggcag tggatctggg
ggttagccat gtgacctcag ggactagcga agtgtccgtc acctagaccc
S  N  R  Y   T  G  V   P  D  R   F  T  G  S   G  S  G>
_________j_________j___VK____j_________j_________>

        560         570         580         590         600
acagatttca ctctcaccat tagcaatatg cagtctgaag acctggctga
tgtctaaagt gagagtggta atcgttatac gtcagacttc tggaccgact
 T  D  F   T  L  T  I   S  N  M   Q  S  E   D  L  A  D>
_________j_________j___VK____j_________j_________>

        610         620         630         640         650
ttttttctgt caacaatata gcaactatcc gtggacgttc ggtggaggca
aaaaaagaca gttgttatat cgttgatagg cacctgcaag ccacctccgt
  F  F  C   Q  Q  Y   S  N  Y   P  W  T   F  G  G  G>
_________j_________j___VK____j_________j_________>

        660         670         680         690         700
ccaagctgga aatcaaagga tctggctcta cttccggtag cggcaaatcc
ggttcgacct ttagtttcct agaccgagat gaaggccatc gccgtttagg
T  K  L  E   I  K>
_______VK_j_______>
                     G   S  G  S   T  S  G  S   G  K  S>
                  ___k_________LINK_____k_________>
```

710 720 730 740 750
tctgaaggca aaggtactag agacgtgaag ctcgtggagt ctgggggagg
agacttccgt ttccatgatc tctgcacttc gagcacctca gacccccctcc
D V K L V E S G G G>
_____h___VH___h_____>
S E G K G T R>
_____LINK_____k_>

760 770 780 790 800
cttagtgaag cttggagggt ccctgaaact ctcctgtgca gcctctggat
gaatcacttc gaacctccca gggactttga gaggacacgt cggagaccta
L V K L G G S L K L S C A A S G>
_____h_____h___VH____h_____h_____>

810 820 830 840 850
tcactttcag taactattac atgtcttggg ttcgccagac tccagagaag
agtgaaagtc attgataatg tacagaaccc aagcggtctg aggtctcttc
F T F S N Y Y M S W V R Q T P E K>
_____h_____h___VH____h_____h_____>

860 870 880 890 900
aggctggagt tggtcgcagc cattaatagt gatggtggta tcacctacta
tccgacctca accagcgtcg gtaattatca ctaccaccat agtggatgat
R L E L V A A I N S D G G I T Y Y>
_____h_____h___VH____h_____h_____>

910 920 930 940 950
tctagacact gtgaagggcc gattcaccat ttcaagagac aatgccaaga
agatctgtga cacttcccgg ctaagtggta aagttctctg ttacggttct
L D T V K G R F T I S R D N A K>
_____h_____h___VH____h_____h_____>

960 970 980 990 1000
acaccctgta cctgcaaatg agcagtctga agtctgagga cacagccttg
tgtgggacat ggacgtttac tcgtcagact tcagactcct gtgtcggaac

```
N  T  L  Y   L  Q  M   S  S  L   K  S  E  D   T  A  L>
_________h_________h___VH____h_________h_________>

      1010        1020        1030        1040        1050
ttttactgtg caagacaccg ctcgggctac ttttctatgg actactgggg
aaaatgacac gttctgtggc gagcccgatg aaaagatacc tgatgacccc
 F  Y  C   A  R  H  R   S  G  Y   F  S  M   D  Y  W  G>
_________h_________h___VH____h_________h_________>

      1060        1070        1080        1090        1100
tcaaggaacc tcagtcaccg tctcctcatg cccaccgtgc ccagcacctg
agttccttgg agtcagtggc agaggagtac gggtggcacg ggtcgtggac
  Q  G  T   S  V  T   V  S  S>
_________h___VH_____h_______>
                                                    A  P>
                                                   _____>
                                          C   P  P  C   P>
                                        __1__HINGE___1___>

      1110        1120        1130        1140        1150
aactcctagg gggaccgtca gtcttcctct tccccccaaa acccaaggac
ttgaggatcc ccctggcagt cagaaggaga agggggggttt tgggttcctg
E  L  L  G   G  P  S   V  F  L   F  P  P  K   P  K  D>
_________i_________i___CH2___i_________i_________>

      1160        1170        1180        1190        1200
accctcatga tctcccggac ccctgaggtc acatgcgtgg tggtggacgt
tgggagtact agagggcctg gggactccag tgtacgcacc accacctgca
 T  L  M   I  S  R  T   P  E  V   T  C  V   V  V  D  V>
_________i_________i___CH2___i_________i_________>

      1210        1220        1230        1240        1250
gagccacgaa gaccctgagg tcaagttcaa ctggtacgtg gacggcgtgg
ctcggtgctt ctgggactcc agttcaagtt gaccatgcac ctgccgcacc
  S  H  E   D  P  E   V  K  F  N   W  Y  V   D  G  V>
```

```
          _____i_____i___CH2___i_____i_____>

               1260       1270       1280       1290       1300
          aggtgcataa tgccaagaca aagccgcggg aggagcagta caacagcacg
          tccacgtatt acggttctgt ttcggcgccc tcctcgtcat gttgtcgtgc
          E  V  H  N  A  K  T  K  P  R  E  E  Q  Y  N  S  T>
          _____i_____i___CH2___i_____i_____>

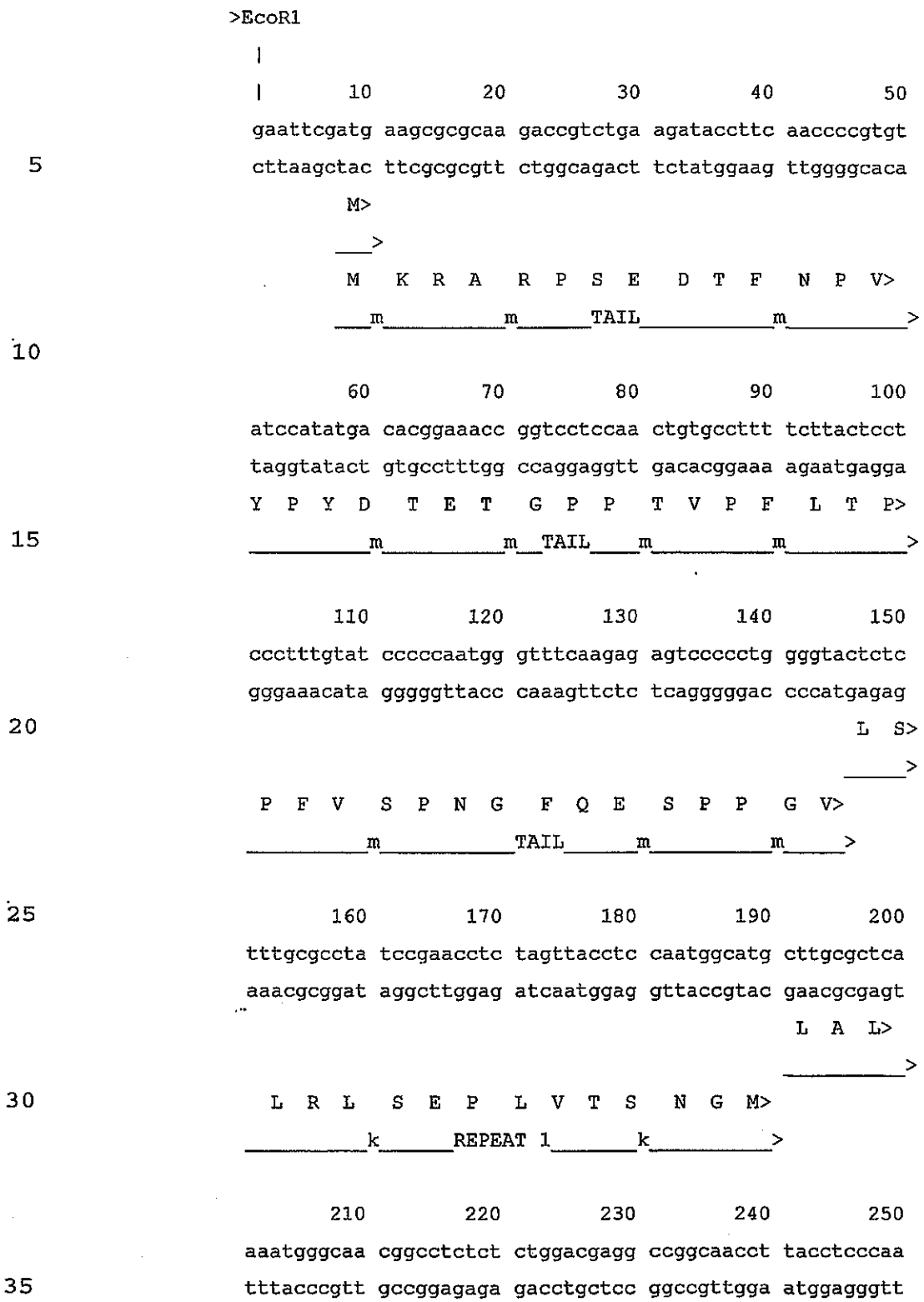
               1310       1320       1330       1340       1350
          taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
          atggcccacc agtcgcagga gtggcaggac gtggtcctga ccgacttacc
          Y  R  V  V  S  V  L  T  V  L  H  Q  D  W  L  N  G>
          _____i_____i___CH2___i_____i_____>

               1360       1370       1380       1390       1400
          caaggagtac aagtgcaagg tctccaacaa agccctccca gccccccatcg
          gttcctcatg ttcacgttcc agaggttgtt tcgggagggt cgggggtagc
            K  E  Y  K  C  K  V  S  N  K  A  L  P  A  P  I>
          _____i_____i___CH2___i_____i_____>

            >XhoI
              |
              |1410
          agtaactcga g
          tcattgagct c
            *  L  E>
           _____a_>
          E>
          __>
```

[SEQ. ID. NO.: 11] - Fiber Construct A7 G250

Sequence Range: 1 to 1702


>EcoR1
10 20 30 40 50
gaattcgatg aagcgcgcaa gaccgtctga gataccttc aaccccgtgt
cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
M>
M K R A R P S E D T F N P V>
m m TAIL m>
60 70 80 90 100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y P Y D T E T G P P T V P F L T P>
m m TAIL m m>
110 120 130 140 150
ccctttgtat cccccaatgg gtttcaagag agtccccctg gggtactctc
gggaaacata gggggttacc caaagttctc tcaggggggac cccatgagag
L S>
P F V S P N G F Q E S P P G V>
m TAIL m m>
160 170 180 190 200
tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca
aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt
L A L>
L R L S E P L V T S N G M>
k REPEAT 1 k>
210 220 230 240 250
aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa
tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt

```
                                                    L   T   S   Q>
                                                   __i_________>
          K   M   G   N   G   L   S   L   D   E   A   G   N>
          _________j_____REPEAT 2_______j________>

          260         270         280         290         300
aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa
ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt
                                                        I   N>
                                                       _____>
          N   V   T   T   V   S   P   P   L   K   K   T   K   S   N>
          _________i_______REPEAT 3 ____i_________i_____>
          310         320         330         340         350
cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg
ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc
                                                    L   T   V>
                                                   _________>
            L   E   I   S   A   P   L   T   V   T   S   E   A>
          _________h_____REPEAT 4______h_________>

          360         370         380         390         400
ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                                L   T   M   Q   S>
                                                ___REPEAT 6_____>
          A   A   A   A   P   L   M   V   A   G   N   T>
          _________g___REPEAT 5_________g_____>

          410         420         430         440         450
caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                            L   S   I   A   T   Q   G>
                                           ______REPEAT 7______>
          Q   A   P   L   T   V   H   D   S   K>
          _________REPEAT 6__f_________>
```

```
          460         470         480         490         500
accctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
                                  L  A>
                                 _____d_>
  P  L  T   V  S  E   G  K>
_________REPEAT 7____e_____>
                                        P  D  V   A  S  L>
                                      ___LUNG SURFACTA____>

          510         520         530         540         550
gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
R  Q  Q  V   E  A  L   Q  G  Q   V  Q  H  L   Q  A  A>
__________LUNG SURFACTANT PROTEIN D TRIM__n_________>

          560         570         580         590         600
tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
                                                 A  K  K  L>
                                               _b_________>
 F  S  Q   Y  K  K  V   E  L  F   P  N  G>
_____LUNG SURFACTANT PROTEIN D TRIM_______>

                                    >Cla1
                                      |
          610         620         630 |       640         650
gaacgacgcc caggccccca agagcgaccc atcgatcgac attgtgatga
cttgctgcgg gtccgggggt tctcgctggg tagctagctg taacactact
  N  D  A   Q  A  P   K  S  D>
__________SPA 48-60__b_______>
                                    P   S  I>
                                   __c_____>
                                          D   I  V  M>
                                       ___q_________>
```

660 670 680 690 700
cccagtctca aagattcatg tccacaacag taggagacag ggtcagcatc
gggtcagagt ttctaagtac aggtgttgtc atcctctgtc ccagtcgtag
T Q S Q R F M S T T V G D R V S I>
_____q_____q___VK____q_____q_____>

710 720 730 740 750
acctgcaagg ccagtcagaa tgtggtttct gctgttgcct ggtatcaaca
tggacgttcc ggtcagtctt acaccaaaga cgacaacgga ccatagttgt
T C K A S Q N V V S A V A W Y Q Q>
_____q_____q___VK____q_____q_____>

760 770 780 790 800
gaaaccagga caatctccta aactactgat ttactcagca tccaatcggt
ctttggtcct gttagaggat ttgatgacta aatgagtcgt aggttagcca
K P G Q S P K L L I Y S A S N R>
_____q_____q___VK____q_____q_____>

810 820 830 840 850
acactggagt ccctgatcgc ttcacaggca gtggatctgg gacagatttc
tgtgacctca gggactagcg aagtgtccgt cacctagacc ctgtctaaag
Y T G V P D R F T G S G S G T D F>
_____q_____q___VK____q_____q_____>

860 870 880 890 900
actctcacca ttagcaatat gcagtctgaa gacctggctg atttttctg
tgagagtggt aatcgttata cgtcagactt ctggaccgac taaaaaagac
T L T I S N M Q S E D L A D F F C>
_____q_____q___VK____q_____q_____>

910 920 930 940 950
tcaacaatat agcaactatc cgtggacgtt cggtggaggc accaagctgg
agttgttata tcgttgatag gcacctgcaa gccacctccg tggttcgacc
Q Q Y S N Y P W T F G G G T K L>
_____q_____q___VK____q_____q_____>

960 970 980 990 1000
aaatcaaagg atctggctct acttccggta gcggcaaatc ctctgaaggc
tttagtttcc tagaccgaga tgaaggccat cgccgtttag gagacttccg
E I K>
_____>
G S G S T S G S G K S S E G>
__r_____r_____LINK_____r_____>

1010 1020 1030 1040 1050
aaaggtacta gagacgtgaa gctcgtggag tctgggggag gcttagtgaa
tttccatgat ctctgcactt cgagcacctc agacccccctc cgaatcactt
D V K L V E S G G G L V K>
_____o_____VH_____o_____>
K G T R>
_____r__>

1060 1070 1080 1090 1100
gcttggaggg tccctgaaac tctcctgtgc agcctctgga ttcactttca
cgaacctccc agggactttg agaggacacg tcggagacct aagtgaaagt
L G G S L K L S C A A S G F T F>
_____o_____o___VH____o_____o_____>

1110 1120 1130 1140 1150
gtaactatta catgtcttgg gttcgccaga ctccagagaa gaggctggag
cattgataat gtacagaacc caagcggtct gaggtctctt ctccgacctc
S N Y Y M S W V R Q T P E K R L E>
_____o_____o___VH____o_____o_____>

1160 1170 1180 1190 1200
ttggtcgcag ccattaatag tgatggtggt atcacctact atctagacac
aaccagcgtc ggtaattatc actaccacca tagtggatga tagatctgtg
L V A A I N S D G G I T Y Y L D T>
_____o_____o___VH____o_____o_____>

1210 1220 1230 1240 1250
tgtgaagggc cgattcacca tttcaagaga caatgccaag aacaccctgt

```
acacttcccg gctaagtggt aaagttctct gttacggttc ttgtgggaca
  V  K  G   R  F  T   I  S  R  D   N  A  K   N  T  L>
_________o_________o___VH____o_________o_________>

     1260       1270       1280       1290       1300
acctgcaaat gagcagtctg aagtctgagg acacagcctt gttttactgt
tggacgttta ctcgtcagac ttcagactcc tgtgtcggaa caaaatgaca
Y  L  Q  M   S  S  L   K  S  E   D  T  A  L   F  Y  C>
_________o_________o___VH____o_________o_________>

     1310       1320       1330       1340       1350
gcaagacacc gctcgggcta cttttctatg gactactggg gtcaaggaac
cgttctgtgg cgagcccgat gaaaagatac ctgatgaccc cagttccttg
 A  R  H   R  S  G  Y   F  S  M   D  Y  W   G  Q  G  T>
_________o_________o___VH____o_________o_________>

     1360       1370       1380       1390       1400
ctcagtcacc gtctcctcat gcccaccgtg cccagcacct gaactcctag
gagtcagtgg cagaggagta cgggtggcac gggtcgtgga cttgaggatc
  S  V  T   V  S  S>
________VH_________>
                                      A  P   E  L  L>
                                     ______CH2_______>
                            C  P  P  C   P>
                          _s___HINGE__s____>

     1410       1420       1430       1440       1450
ggggaccgtc agtcttcctc ttcccccaa aacccaagga caccctcatg
cccctggcag tcagaaggag aaggggggtt ttgggttcct gtgggagtac
G  G  P  S   V  F  L   F  P  P   K  P  K  D   T  L  M>
_________p_________p___CH2___p_________p_________>

     1460       1470       1480       1490       1500
atctcccgga cccctgaggt cacatgcgtg gtggtggacg tgagccacga
tagagggcct ggggactcca gtgtacgcac caccacctgc actcggtgct
```

```
                I  S  R   T  P  E  V   T  C  V   V  V  D   V  S  H  E>
               _____p_____p___CH2___p_____p_____>

                   1510       1520       1530       1540       1550
5              agaccctgag gtcaagttca actggtacgt ggacggcgtg gaggtgcata
               tctgggactc cagttcaagt tgaccatgca cctgccgcac ctccacgtat
                 D  P  E   V  K  F   N  W  Y  V   D  G  V   E  V  H>
               _____p_____p___CH2___p_____p_____>

10                 1560       1570       1580       1590       1600
               atgccaagac aaagccgcgg gaggagcagt acaacagcac gtaccgggtg
               tacggttctg tttcggcgcc ctcctcgtca tgttgtcgtg catggcccac
               N  A  K  T   K  P  R   E  E  Q   Y  N  S  T   Y  R  V>
               _____p_____p___CH2___p_____p_____>
15
                   1610       1620       1630       1640       1650
               gtcagcgtcc tcaccgtcct gcaccaggac tggctgaatg gcaaggagta
               cagtcgcagg agtggcagga cgtggtcctg accgacttac cgttcctcat
                V  S  V   L  T  V  L   H  Q  D   W  L  N   G  K  E  Y>
20             _____p_____p___CH2___p_____p_____>

                                                          >Xho1
                                                           |
                   1660       1670       1680       1690       1700
25             caagtgcaag gtctccaaca aagccctccc agcccccatc gagtaactcg
               gttcacgttc cagaggttgt ttcgggaggg tcgggggtag ctcattgagc
                                                        *  L>
                                                       _____>
                 K  C  K   V  S  N   K  A  L  P   A  P  I   E>
30             _____p_____CH2_____p_____p___>

               ag
               tc
               E>
35             __>
```

[SEQ. ID. NO.: 12] - Fiber Construct A1 C242

Sequence Range: 1 to 1111

```
>EcoR1
 |
 |    10         20         30         40         50
gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
       M>
      ___>
       M   K   R   A   R   P   S   E   D   T   F   N   P   V>
      ___o_________o_____TAIL__________o_________>

       60         70         80         90         100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y  P   Y   D   T   E   T   G   P   P   T   V   P   F   L   T   P>
_________o_________o__TAIL___o_________o_________>

      110        120        130        140        150
ccctttgtat cccccaatgg gtttcaagag agtcccccctg gggtactctc
gggaaacata gggggttacc caaagttctc tcaggggggac cccatgagag
                                                    L   S>
                                                   _____>
 P   F   V   S   P   N   G   F   Q   E   S   P   P   G   V>
_________o_________TAIL_____o_________o_____>

      160        170        180        190        200
tttgcgccta tccgaacctc tagttacctc caatggcatg cctgacgtag
aaacgcggat aggcttggag atcaatggag gttaccgtac ggactgcatc
  L   R   L   S   E   P   L   V   T   S   N   G   M>
_________m_____REPEAT 1_____m_________>
```

```
                                                                P  D  V>
                                                               _________>

          210         220         230         240         250
caagcttacg acaacaggta gaagccttgc aagggcaggt acaacactta
gttcgaatgc tgttgtccat cttcggaacg ttcccgtcca tgttgtgaat
A  S  L  R   Q  Q  V   E  A  L   Q  G  Q  V   Q  H  L>
__________LUNG SURFACTANT PROTEIN D TRIM__p_________>

          260         270         280         290         300
caggcggcat ttagccaata caaaaaggta gagttgtttc caaacggagc
gtccgccgta aatcggttat gtttttccat ctcaacaaag gtttgcctcg
                                                              A>
                                                             __>
 Q  A  A   F  S  Q  Y   K  K  V   E  L  F   P  N  G>
_________LUNG SURFACTANT PROTEIN D TRIM___p_______>

                                                    >Cla1
                                                      |
          310         320         330         340         350
caagaagctg aacgacgccc aggccccccaa gagcgaccca tcgatcgata
gttcttcgac ttgctgcggg tccgggggtt ctcgctgggt agctagctat
  K  K  L   N  D  A   Q  A  P  K   S  D>
__________b____SPA 48-60_______b_______>
                                                     P   S  I>
                                                   ___c_____>
                                                            D>
                                                           ____>

          360         370         380         390         400
ttgtgatgac tcaggctgca ccctctgtac ctgtcactcc tggagagtca
aacactactg agtccgacgt gggagacatg gacagtgagg acctctcagt
I  V  M  T   Q  A  A   P  S  V   P  V  T  P   G  E  S>
_________e_________VK PHARMACIA__________e_________>
```

```
          410        420        430        440        450
gtatccatct cctgcaggtc tagtaagagt ctcctgcata gtaatggcaa
cataggtaga ggacgtccag atcattctca gaggacgtat cattaccgtt
 V  S  I   S  C  R  S   S  K  S   L  L  H   S  N  G  N>
_________e________VK PHARMACIA_________e________>
                                S   L  L  H   S  N  G  N>
                               ___g_____CDR1_________>

          460        470        480        490        500
cacttacttg tattggttcc tgcagaggcc aggccagtct cctcagctcc
gtgaatgaac ataaccaagg acgtctccgg tccggtcaga ggagtcgagg
  T  Y  L   Y  W  F   L  Q  R  P   G  Q  S   P  Q  L>
_________e________VK PHARMACIA_________e________>
  T  Y  L   Y>
_________g___>

          510        520        530        540        550
tgatatatcg gatgtccaac cttgtctcag gagtcccaga caggttcagt
actatatagc ctacaggttg gaacagagtc ctcagggtct gtccaagtca
L  I  Y  R   M  S  N   L  V  S   G  V  P  D   R  F  S>
_________e________VK PHARMACIA_________e________>
          R   M  S  N   L  V  S>
         __f_____CDR2_________>

          560        570        580        590        600
ggcagtgggt caggaactgc tttcacactg agaatcagta gagtggaggc
ccgtcaccca gtccttgacg aaagtgtgac tcttagtcat ctcacctccg
 G  S  G   S  G  T  A   F  T  L   R  I  S   R  V  E  A>
_________e________VK PHARMACIA_________e________>

          610        620        630        640        650
tgaggatgtg ggtgtttatt actgtctgca acatctagag tatccgttca
actcctacac ccacaaataa tgacagacgt tgtagatctc ataggcaagt
                                L  Q   H  L  E   Y  P  F>
                              _____d____CDR3_d_________>
```

```
                                E  D  V   G  V  Y   Y  C  L  Q   H  L  E   Y  P  F>
          _____e_____VK PHARMACIA_____e_____>

                660        670        680        690        700
cgttcggtcc tgggaccaag ctggagctga aacggccccc ggactttgtt
gcaagccagg accctggttc gacctcgact ttgccggggg cctgaaacaa
T>
__>
T  F  G  P   G  T  K   L  E  L   K  R>
_____e_VK PHARMACIA_____e_____>
                                                P  P   D  F  V>
                                              _____825_____>

                710        720        730        740        750
ccccccggccg ctagtttccc tgatcactcc cctcgtggcc aggtccagtt
gggggccggc gatcaaaggg actagtgagg ggagcaccgg tccaggtcaa
                                                      Q  V  Q  L>
                                                    _i_____>
 P  P  A   A  S  F  P   D  H  S   P  R  G>
_____l_____825_____l_____>

                760        770        780        790        800
ggtgcagtct ggacctgagc tgaagaagcc tggagagaca gtcaagatct
ccacgtcaga cctggactcg acttcttcgg acctctctgt cagttctaga
  V  Q  S   G  P  E   L  K  K  P   G  E  T   V  K  I>
_____i_____VH PHARMACIA [SPLIT]_____i_____>
                810        820        830        840        850
cctgcaaggc ttctgattat accttcacat actatggaat gaactgggtg
ggacgttccg aagactaata tggaagtgta tgatacctta cttgacccac
S  C  K  A   S  D  Y   T  F  T   Y  Y  G  M   N  W  V>
_____i_____VH PHARMACIA [SPLIT]_____i_____>
                                                Y  Y  G  M   N>
                                              _k____CDR1__k____>

                860        870        880        890        900
```

```
aagcaggctc cgggaaaggg tttaaagtgg atgggctgga tagacaccac
ttcgtccgag gcccttccc aaatttcacc tacccgacct atctgtggtg
 K  Q  A   P  G  K  G   L  K  W   M  G  W   I  D  T  T>
_________i_____VH PHARMACIA [SPLIT]_______i_________>
                                          W   I  D  T  T>
                                       ___CDR2 [SP____>

      910        920        930        940        950
cactggagag ccaacatatg ctgaagattt taagggacgg attgccttct
gtgacctctc ggttgtatac gacttctaaa attccctgcc taacggaaga
  T  G  E   P  T  Y   A  E  D  F   K  G  R   I  A  F>
_________i_____VH PHARMACIA [SPLIT]_______i_________>
  T  G  E   P  T  Y   A  E  D  F   K  G>
_________j__CDR2 [SPLIT]_______j_______>

      960        970        980        990       1000
ctttggagac ctctgccagc actgcctatt tgcagatcaa aaacctcaaa
gaaacctctg gagacggtcg tgacggataa acgtctagtt tttggagttt
S  L  E  T   S  A  S   T  A  Y   L  Q  I   K  N  L  K>
_________i_____VH PHARMACIA [SPLIT]_______i_________>

     1010       1020       1030       1040       1050
aatgaggaca cggctacata tttctgtgca agacgggggc cttacaactg
ttactcctgt gccgatgtat aaagacacgt tctgcccccg gaatgttgac
                                       R  G   P  Y  N  W>
                                     _____CDR3_______>
 N  E  D   T  A  T  Y   F  C  A   R  R  G   P  Y  N  W>
_________i_____VH PHARMACIA [SPLIT]_____i_________>
     1060       1070       1080       1090       1100
gtactttgat gtctggggcc aagggaccac ggtcaccgtc tcctcactcg
catgaaacta cagaccccgg ttccctggtg ccagtggcag aggagtgagc
                                                    L>
                                                   ____>
  Y  F  D   V>
_________h___>
```

```
  Y  F  D   V  W  G   Q  G  T  T   V  T  V   S  S>
_________i___VH PHARMACIA [SPLIT]________i_____>

   >Xho1
     |
     |1110
attaactcga g
taattgagct c
D  *  L  E>
_________a_>
```

[SEQ. ID. NO.: 13] - Fiber Construct A7 C242

Sequence Range: 1 to 1402

```
>EcoR1
  |
  |      10         20         30         40         50
gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
       M>
      ___>
       M   K  R  A   R  P  S  E   D  T  F   N  P  V>
      ___v_________v_____TAIL_________v_________>

       60         70         80         90         100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y  P  Y  D   T  E  T   G  P  P   T  V  P  F   L  T  P>
_________v_________v__TAIL___v_________v_________>

      110        120        130        140        150
ccctttgtat cccccaatgg gtttcaagag agtccccctg gggtactctc
```

gggaaacata gggggttacc caaagttctc tcaggggac cccatgagag

L S>

_____>

P F V S P N G F Q E S P P G V>

_____v_____TAIL_____v_____v_____>

160 170 180 190 200

tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca

aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt

L A L>

_____>

L R L S E P L V T S N G M>

_____t_____REPEAT 1_____t_____>

210 220 230 240 250

aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa

tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt

L T S Q>

__r_____>

K M G N G L S L D E A G N>

_____s_____REPEAT 2_____s_____>

260 270 280 290 300

aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa

ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt

I N>

_____>

N V T T V S P P L K K T K S N>

_____r_____REPEAT 3 ____r_____r_____>

310 320 330 340 350

cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg

ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc

L T V>

_____>

```
                L  E  I   S  A  P   L  T  V  T   S  E  A>
          _____q_____REPEAT 4_____q_____>

          360        370        380        390        400
ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                          L  T   M  Q  S>
                                        ___REPEAT 6_____>
A  A  A  A   P  L  M   V  A  G   N  T>
_____p___REPEAT 5_____p_____>

          410        420        430        440        450
caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                    L  S  I   A  T  Q  G>
                                   ______REPEAT 7_______>
 Q  A  P   L  T  V  H   D  S  K>
_____REPEAT 6__o_____>

          460        470        480        490        500
acccctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
                               L  A>
                             _____m_>
  P  L  T   V  S  E   G  K>
_____REPEAT 7____n_____>
                                    P  D  V   A  S  L>
                                  ___LUNG SURFACTA____>

          510        520        530        540        550
gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
R  Q  Q  V   E  A  L   Q  G  Q   V  Q  H  L   Q  A  A>
_____LUNG SURFACTANT PROTEIN D TRIM__w_____>
          560        570        580        590        600
```

```
tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
                                                A  K  K  L>
                                              _b_________>
 F  S  Q   Y  K  K  V   E  L  F   P  N  G>
_____LUNG SURFACTANT PROTEIN D TRIM_______>

                                    >Cla1
                                      |
      610         620         630   |     640         650
gaacgacgcc caggccccca agagcgaccc atcgatcgat attgtgatga
cttgctgcgg gtccgggggt tctcgctggg tagctagcta taacactact
  N  D  A   Q  A  P   K  S  D>
__________SPA 48-60__b________>
                                    P   S  I>
                                   __c_____>
                                          D   I  V  M>
                                        ___e_________>

       660         670         680         690         700
ctcaggctgc accctctgta cctgtcactc ctggagagtc agtatccatc
gagtccgacg tgggagacat ggacagtgag gacctctcag tcataggtag
T  Q  A  A   P  S  V   P  V  T   P  G  E  S   V  S  I>
__________e_________VK PHARMACIA__________e_________>

       710         720         730         740         750
tcctgcaggt ctagtaagag tctcctgcat agtaatggca acacttactt
aggacgtcca gatcattctc agaggacgta tcattaccgt tgtgaatgaa
 S  C  R   S  S  K  S   L  L  H   S  N  G   N  T  Y  L>
__________e_________VK PHARMACIA__________e_________>
                      S   L  L  H   S  N  G   N  T  Y  L>
                    __g_________g_CDR1_____g_________>

       760         770         780         790         800
gtattggttc ctgcagaggc caggccagtc tcctcagctc ctgatatatc
```

cataaccaag gacgtctccg gtccggtcag aggagtcgag gactatatag
Y W F L Q R P G Q S P Q L L I Y>
_____e_____VK PHARMACIA_____e_____>
_>
Y>
____>

810 820 830 840 850
ggatgtccaa ccttgtctca ggagtcccag acaggttcag tggcagtggg
cctacaggtt ggaacagagt cctcagggtc tgtccaagtc accgtcaccc
R M S N L V S G V P D R F S G S G>
_____e_____VK PHARMACIA_____e_____>
R M S N L V S>
_____CDR2_____>

860 870 880 890 900
tcaggaactg ctttcacact gagaatcagt agagtggagg ctgaggatgt
agtccttgac gaaagtgtga ctcttagtca tctcacctcc gactcctaca
S G T A F T L R I S R V E A E D V>
_____e_____VK PHARMACIA_____e_____>

910 920 930 940 950
gggtgtttat tactgtctgc aacatctaga gtatccgttc acgttcggtc
cccacaaata atgacagacg ttgtagatct cataggcaag tgcaagccag
L Q H L E Y P F T>
____d_____CDR3_____d___>
G V Y Y C L Q H L E Y P F T F G>
_____e_____VK PHARMACIA_____e_____>

960 970 980 990 1000
ctgggaccaa gctggagctg aaacggcccc cggactttgt tcccccggcc
gaccctggtt cgacctcgac tttgccgggg gcctgaaaca agggggccgg
P G T K L E L K R>
_____VK PHARMACIA__e_____>
P P D F V P P A>

```
                                                  ____l_____825_l_________>

      1010       1020       1030       1040       1050
gctagtttcc ctgatcactc ccctcgtggc caggtccagt tggtgcagtc
cgatcaaagg gactagtgag gggagcaccg gtccaggtca accacgtcag
                                          Q  V  Q   L  V  Q  S>
                                         ___VH PHARMACIA [____>
 A  S  F   P  D  H  S   P  R  G>
_____l___825____l_________>

      1060       1070       1080       1090       1100
tggacctgag ctgaagaagc ctggagagac agtcaagatc tcctgcaagg
acctggactc gacttcttcg gacctctctg tcagttctag aggacgttcc
  G  P  E   L  K  K   P  G  E  T  V  K  I   S  C  K>
_________i_____VH PHARMACIA [SPLIT]_______i_________>

      1110       1120       1130       1140       1150
cttctgatta taccttcaca tactatggaa tgaactgggt gaagcaggct
gaagactaat atggaagtgt atgatacctt acttgaccca cttcgtccga
A  S  D  Y   T  F  T   Y  Y  G   M  N  W  V   K  Q  A>
_________i_____VH PHARMACIA [SPLIT]_______i_________>
                              Y  Y  G   M  N>
                             _____CDR1_k_____>

      1160       1170       1180       1190       1200
ccgggaaagg gtttaaagtg gatgggctgg atagacacca ccactggaga
ggccctttcc caaatttcac ctacccgacc tatctgtggt ggtgacctct
 P  G  K   G  L  K  W   M  G  W   I  D  T   T  T  G  E>
_________i_____VH PHARMACIA [SPLIT]_______i_________>
                                    W   I  D  T   T  T  G  E>
                                   ___j__CDR2 [SPLIT]_______>

      1210       1220       1230       1240       1250
gccaacatat gctgaagatt ttaagggacg gattgccttc tctttggaga
cggttgtata cgacttctaa aattccctgc ctaacggaag agaaacctct
```

```
  P  T  Y  A  E  D  F  K  G  R  I  A  F  S  L  E>
_____i____VH PHARMACIA [SPLIT]_____i_____>
  P  T  Y  A  E  D  F  K  G>
_____CDR2 [SPLIT]_j_____>

     1260       1270       1280       1290       1300
cctctgccag cactgcctat ttgcagatca aaaacctcaa aaatgaggac
ggagacggtc gtgacggata aacgtctagt ttttggagtt tttactcctg
T  S  A  S  T  A  Y  L  Q  I  K  N  L  K  N  E  D>
_____i____VH PHARMACIA [SPLIT]_____i_____>

     1310       1320       1330       1340       1350
acggctacat atttctgtgc aagacggggg ccttacaact ggtactttga
tgccgatgta taaagacacg ttctgccccc ggaatgttga ccatgaaact
                              R  G  P  Y  N  W  Y  F  D>
                           _____h___CDR3__h_____>
 T  A  T  Y  F  C  A  R  R  G  P  Y  N  W  Y  F  D>
_____i____VH PHARMACIA [SPLIT]_____i_____>

                                                >Xho1
                                                 |
     1360       1370       1380       1390       1400
tgtctggggc caagggacca cggtcaccgt ctcctcactc gattaactcg
acagaccccg gttccctggt gccagtggca gaggagtgag ctaattgagc
                                          L  D  *  L>
                                        ___a_____>

  V>
____>
  V  W  G  Q  G  T  T  V  T  V  S  S>
_____VH PHARMACIA [SPLIT]___i_____>

ag
tc
E>
__>
```

[SEQ. ID. NO.: 14] -

Sequence Range: 1 to 174 - Affibody ZIgG

```
          10         20         30         40         50
gtagacaaca aattcaacaa agaacaacaa aacgcgttct atgagatctt
catctgttgt ttaagttgtt tcttgttgtt ttgcgcaaga tactctagaa
 V  D  N   K  F  N  K   E  Q  Q   N  A  F   Y  E  I  L>
_________a_________a_Z-DOM?N__a_________a_________>

          60         70         80         90        100
acatttacct aacttaaacg aagaacaacg aaacgccttc atccaaagtt
tgtaaatgga ttgaatttgc ttcttgttgc tttgcggaag taggtttcaa
  H  L  P   N  L  N   E  E  Q  R   N  A  F   I  Q  S>
_________a_________a_Z-DOM?N__a_________a_________>

         110        120        130        140        150
taaaagatga cccaagccaa agcgctaact tgctagcaga agctaaaaag
attttctact gggttcggtt tcgcgattga acgatcgtct tcgatttttc
L  K  D  D  P  S  Q   S  A  N   L  L  A  E   A  K  K>
_________a_________a_Z-DOM?N__a_________a_________>

          160        170
ctaaatgatg ctcaggcgcc gaaa
gatttactac gagtccgcgg cttt
 L  N  D   A  Q  A  P   K>
_________Z-DOM?N____a___>
```

[SEQ. ID. NO.: 15] - Affibody ZIgA

Sequence Range: 1 to 174

```
          10         20         30         40         50
gtagacaaca aattcaacaa agaaacaata caagcgagtc aagagatcag
catctgttgt ttaagttgtt tctttgttat gttcgctcag ttctctagtc
 V  D  N   K  F  N  K   E  T  I   Q  A  S   Q  E  I  R>
_________a__________AFFI IGA _a_________a_________>

          60         70         80         90        100
actattacct aacttaaacg gtagacaaaa gcttgccttc atccacagtt
tgataatgga ttgaatttgc catctgtttt cgaacggaag taggtgtcaa
  L  L  P   N  L  N   G  R  Q  K   L  A  F   I  H  S>
_________a__________AFFI IGA _a_________a_________>

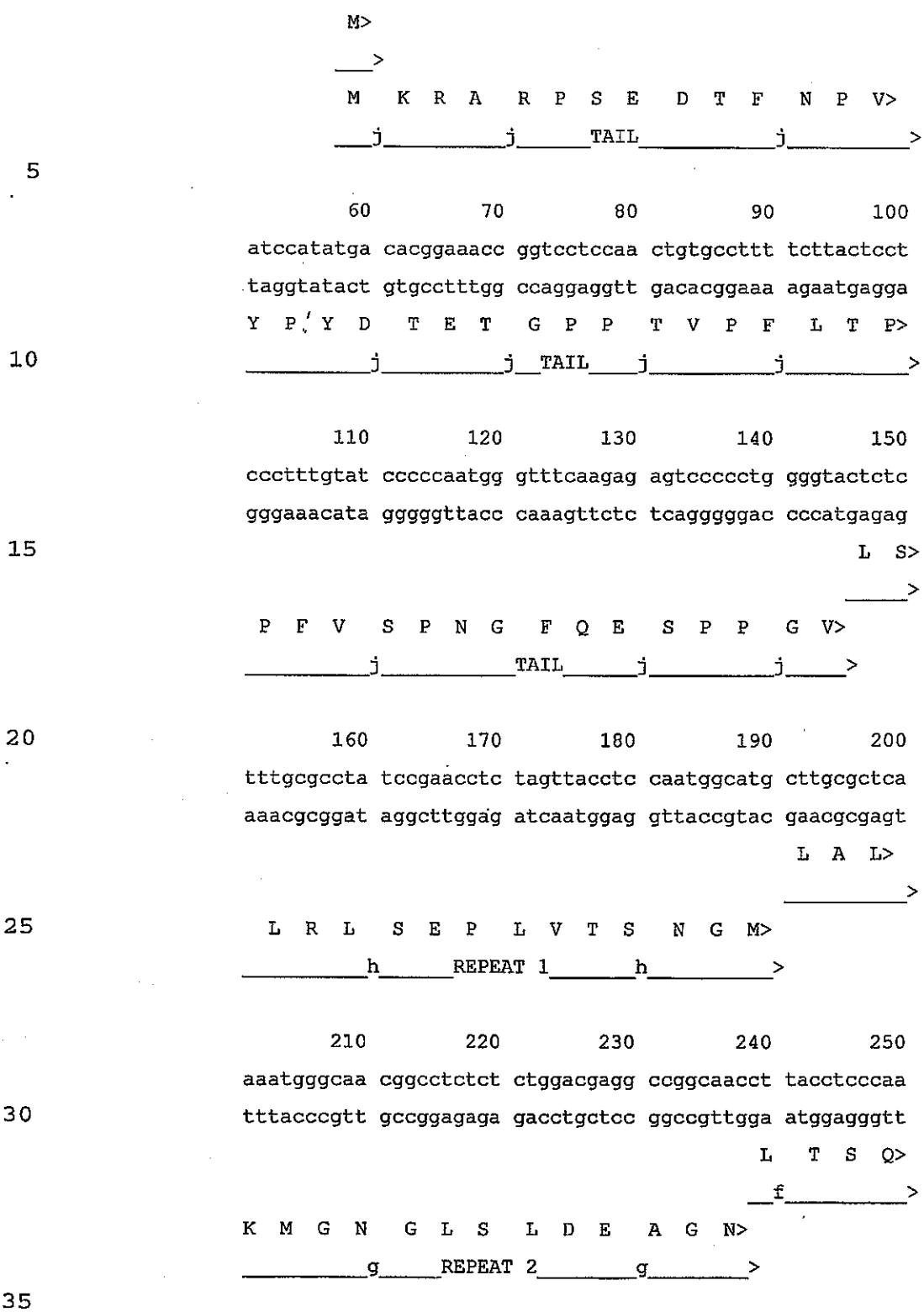
        110        120        130        140        150
tacttgatga cccaagccaa agcgctaact tgctagcaga agctaaaaag
atgaactact gggttcggtt tcgcgattga acgatcgtct tcgatttttc
L  L  D  D   P  S  Q   S  A  N   L  L  A  E   A  K  K>
_________a__________AFFI IGA _a_________a_________>

        160        170
ctaaatgatg ctcaggcgcc gaaa
gatttactac gagtccgcgg cttt
 L  N  D   A  Q  A  P   K>
________AFFI IGA ____a____>
```

[SEQ. ID. NO.: 16] - Recombinant Fiber A7 ZIgG

Sequence Range: 1 to 820

```
>EcoR1
 |
 |        10         20         30         40         50
gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
```


M>
M K R A R P S E D T F N P V>
TAIL
60 70 80 90 100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y P Y D T E T G P P T V P F L T P>
TAIL
110 120 130 140 150
ccctttgtat cccccaatgg gtttcaagag agtcccccctg gggtactctc
gggaaacata gggggttacc caaagttctc tcaggggggac cccatgagag
L S>
P F V S P N G F Q E S P P G V>
TAIL
160 170 180 190 200
tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca
aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt
L A L>
L R L S E P L V T S N G M>
REPEAT 1
210 220 230 240 250
aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa
tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt
L T S Q>
K M G N G L S L D E A G N>
REPEAT 2

```
          260         270         280         290         300
aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa
ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt
                                                  I  N>
                                                  _____>
 N  V  T   T  V  S  P   P  L  K   K  T  K   S  N>
_________f_______REPEAT 3 ____f_________f_____>

          310         320         330         340         350
cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg
ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc
                                                L  T  V>
                                                _________>
  L  E  I   S  A  P   L  T  V  T   S  E  A>
_________e_____REPEAT 4_______e_________>

          360         370         380         390         400
ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                          L  T   M  Q  S>
                                          ___REPEAT 6_____>
A  A  A  A   P  L  M   V  A  G   N  T>
_________d___REPEAT 5_________d_____>

          410         420         430         440         450
caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                    L  S  I   A  T  Q  G>
                                    ______REPEAT 7_______>
 Q  A  P   L  T  V  H   D  S  K>
__________REPEAT 6__c__________>

          460         470         480         490         500
acccctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
```

```
                                                    L  A>
                                                  _____a_>
                P  L  T   V  S  E   G  K>
          _________REPEAT 7____b_____>
                                                       P  D  V   A  S  L>
                                                    ___LUNG SURFACTA____>

          510         520         530         540         550
gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
R  Q  Q  V   E  A  L   Q  G  Q   V  Q  H  L   Q  A  A>
__________LUNG SURFACTANT PROTEIN D TRIM__k__________>

          560         570         580         590         600
tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
 F  S  Q   Y  K  K  V   E  L  F   P  N  G>
_____LUNG SURFACTANT PROTEIN D TRIM_______>
                                                          A  K  K  L>
                                                       _l_________>

                                          >Cla1
                                            |
          610         620         630       |   640         650
gaacgacgcc caggccccca agagcgaccc atcgatcgta gacaacaaat
cttgctgcgg gtccgggggt tctcgctggg tagctagcat ctgttgttta
  N  D  A   Q  A  P   K  S  D>
__________SPA 48-60__1_______>
                                       P   S  I>
                                    __m_____>
                                                 V   D  N  K>
                                              ___o_________>

          660         670         680         690         700
tcaacaaaga acaacaaaac gcgttctatg agatcttaca tttacctaac
```

```
agttgtttct tgttgttttg cgcaagatac tctagaatgt aaatggattg
F  N  K  E   Q  Q  N   A  F  Y   E  I  L  H   L  P  N>
_________o_________o_Z-DOM?N__o_________o_________>

      710        720        730        740        750
ttaaacgaag aacaacgaaa cgccttcatc caaagtttaa aagatgaccc
aatttgcttc ttgttgcttt gcggaagtag gtttcaaatt ttctactggg
 L  N  E   E  Q  R  N   A  F  I   Q  S  L   K  D  D  P>
_________o_________o_Z-DOM?N__o_________o_________>

      760        770        780        790        800
aagccaaagc gctaacttgc tagcagaagc taaaaagcta aatgatgctc
ttcggtttcg cgattgaacg atcgtcttcg atttttcgat ttactacgag
  S  Q  S   A  N  L   L  A  E  A   K  K  L   N  D  A>
_________o_________o_Z-DOM?N__o_________o_________>
            >Xho1
              |
      810     |  820
aggcgccgaa ataactcgag
tccgcggctt tattgagctc
            *>
           ___>
Q  A  P  K>
_________o_>
```

[SEQ. ID. NO.: 17] - Recombinant Fiber A7 ZIgA

Sequence Range: 1 to 820

```
          >EcoR1
            |
            |      10        20        30        40        50
```

```
gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
       M>
      ___>
5      M   K  R  A   R  P  S  E   D  T   F   N  P  V>
      ___j_________j_____TAIL_________j_________>

       60         70         80         90         100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
10 taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y  P  Y  D   T  E  T   G  P  P   T  V  P  F   L  T  P>
_________j_________j__TAIL___j_________j_________>

       110        120        130        140        150
15 ccctttgtat cccccaatgg gtttcaagag agtcccccctg gggtactctc
gggaaacata gggggttacc caaagttctc tcaggggggac cccatgagag
                                                  L  S>
                                                 _____>
 P  F  V   S  P  N  G   F  Q  E   S  P  P   G  V>
20 _________j_________TAIL_____j_________j_____>

       160        170        180        190        200
tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca
aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt
25                                             L  A  L>
                                              _________>
  L  R  L   S  E  P   L  V  T  S   N  G  M>
_________h_____REPEAT 1_____h_________>

30     210        220        230        240        250
aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa
tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt
                                          L   T  S  Q>
                                         __f_________>
35 K  M  G  N   G  L  S   L  D  E   A  G  N>
```

```
          _________g_____REPEAT 2________g________>

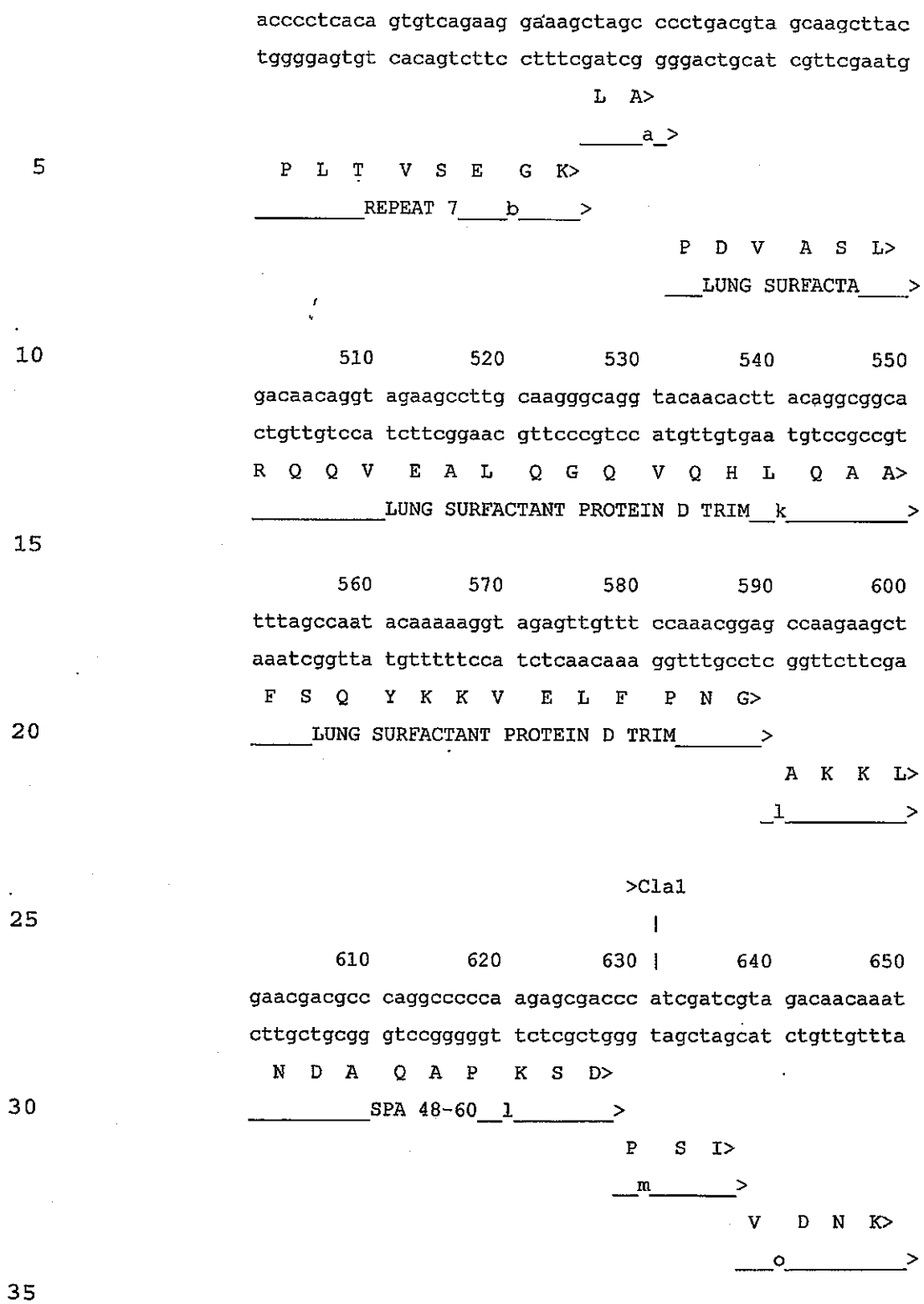
                260        270        280        290        300
          aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa
          ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt
                                                          I  N>
                                                         _____>
           N  V  T   T  V  S  P   P  L  K   K  T  K   S  N>
          _________f_______REPEAT 3 ___f_________f____>

                310        320        330        340        350
          cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg
          ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc
                                                    L  T  V>
                                                   _________>
            L  E  I   S  A  P   L  T  V  T   S  E  A>
          _________e_____REPEAT 4_______e_________>

                360        370        380        390        400
          ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
          gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                                 L  T   M  Q  S>
                                                ___REPEAT 6_____>
          A  A  A  A   P  L  M   V  A  G   N  T>
          _________d___REPEAT 5_________d_____>

                410        420        430        440        450
          caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
          gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                           L  S  I   A  T  Q  G>
                                          ______REPEAT 7_______>
           Q  A  P   L  T  V  H   D  S  K>
          _________REPEAT 6__c_________>

                460        470        480        490        500
```


acccctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
L A>
a
P L T V S E G K>
REPEAT 7 b
P D V A S L>
LUNG SURFACTA
510 520 530 540 550
gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
R Q Q V E A L Q G Q V Q H L Q A A>
LUNG SURFACTANT PROTEIN D TRIM k
560 570 580 590 600
tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
F S Q Y K K V E L F P N G>
LUNG SURFACTANT PROTEIN D TRIM
A K K L>
l
>ClaI
610 620 630 640 650
gaacgacgcc caggcccca agagcgaccc atcgatcgta gacaacaaat
cttgctgcgg gtccgggggt tctcgctggg tagctagcat ctgttgttta
N D A Q A P K S D>
SPA 48-60 l
P S I>
m
V D N K>
o

```
       660         670         680         690         700
tcaacaaaga aacaatacaa gcgagtcaag agatcagact attacctaac
agttgtttct ttgttatgtt cgctcagttc tctagtctga taatggattg
F  N  K  E  T  I  Q  A  S  Q  E  I  R  L  L  P  N>
_________o__________AFFI IGA _o_________o_________>

       710         720         730         740         750
ttaaacggta gacaaaagct tgccttcatc cacagtttac ttgatgaccc
aatttgccat ctgttttcga acggaagtag gtgtcaaatg aactactggg
 L  N  G  R  Q  K  L  A  F  I  H  S  L  L  D  D  P>
_________o__________AFFI IGA _o_________o_________>

       760         770         780         790         800
aagccaaagc gctaacttgc tagcagaagc taaaaagcta aatgatgctc
ttcggtttcg cgattgaacg atcgtcttcg atttttcgat ttactacgag
  S  Q  S  A  N  L  L  A  E  A  K  K  L  N  D  A>
_________o__________AFFI IGA _o_________o_________>

                >Xho1
                  |
       810        | 820
aggcgccgaa ataactcgag
tccgcggctt tattgagctc
                *>
                ___>
Q  A  P  K>
_________o_>
```

[SEQ. ID. NO.: 18] - Recombinant Fiber A7 ZIgG/ZIgA

Sequence Range: 1 to 1003

```
>EcoR1
 |
 |      10         20         30         40         50
 gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
 cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
        M>
       ___>
        M   K  R  A   R  P  S  E   D  T  F   N  P  V>
      ___j_________j_____TAIL_________j_________>

        60         70         80         90        100
 atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
 taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
 Y  P  Y  D   T  E  T   G  P  P   T  V  P  F   L  T  P>
 _________j_________j__TAIL___j_________j_________>

       110        120        130        140        150
 ccctttgtat cccccaatgg gtttcaagag agtccccctg gggtactctc
 gggaaacata gggggttacc caaagttctc tcagggggac cccatgagag
                                                  L  S>
                                                 _____>
  P  F  V   S  P  N  G   F  Q  E   S  P  P   G  V>
 _________j_________TAIL_____j_________j_____>

       160        170        180        190        200
 tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca
 aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt
                                              L  A  L>
                                             _________>
   L  R  L   S  E  P   L  V  T  S   N  G  M>
 _________h_____REPEAT 1_____h_________>

       210        220        230        240        250
 aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa
 tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt
```

```
                                                        L   T   S   Q>
                                                      __f_________>
                K   M   G   N   G   L   S   L   D   E   A   G   N>
                _________g_____REPEAT 2_______g_______>

        260         270         280         290         300
aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa
ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt
                                                            I   N>
                                                          _____>
 N   V   T   T   V   S   P   P   L   K   K   T   K   S   N>
_________f_______REPEAT 3 ____f_________f_____>

        310         320         330         340         350
cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg
ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc
                                                        L   T   V>
                                                      _________>
  L   E   I   S   A   P   L   T   V   T   S   E   A>
_________e_____REPEAT 4______e_________>

        360         370         380         390         400
ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                                    L   T   M   Q   S>
                                                  ___REPEAT 6_____>
A   A   A   A   P   L   M   V   A   G   N   T>
_________d___REPEAT 5________d_____>

        410         420         430         440         450
caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                            L   S   I   A   T   Q   G>
                                          ______REPEAT 7_______>
 Q   A   P   L   T   V   H   D   S   K>
```

```
          _____REPEAT 6__c_____>

                460        470        480        490        500
          acccctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
          tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
                                        L  A>
                                      _____a_>
            P  L  T   V  S  E   G  K>
          _____REPEAT 7___b_____>
                                              P  D  V   A  S  L>
                                            ___LUNG SURFACTA____>

                510        520        530        540        550
          gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
          ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
          R  Q  Q  V   E  A  L   Q  G  Q   V  Q  H  L   Q  A  A>
          _____LUNG SURFACTANT PROTEIN D TRIM__k_____>

                560        570        580        590        600
          tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
          aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
           F  S  Q   Y  K  K  V   E  L  F   P  N  G>
          _____LUNG SURFACTANT PROTEIN D TRIM_____>
                                                    A  K  K  L>
                                                  _l_____>

                                        >Cla1
                                          |
                610        620        630 |      640        650
          gaacgacgcc caggccccca agagcgaccc atcgatcgta gacaacaaat
          cttgctgcgg gtccgggggt tctcgctggg tagctagcat ctgttgttta
            N  D  A   Q  A  P   K  S  D>
          _____SPA 48-60__l_____>
                                        P   S  I>
                                      __m_____>
```

```
                                                    V   D   N   K>
                                                   ___o_________>

          660         670         680         690         700
tcaacaaaga acaacaaaac gcgttctatg agatcttaca tttacctaac
agttgtttct tgttgttttg cgcaagatac tctagaatgt aaatggattg
F   N   K   E   Q   Q   N   A   F   Y   E   I   L   H   L   P   N>
_________o_________o_Z-DOM?N__o_________o_________>

          710         720         730         740         750
ttaaacgaag aacaacgaaa cgccttcatc caaagtttaa aagatgaccc
aatttgcttc ttgttgcttt gcggaagtag gtttcaaatt ttctactggg
 L   N   E   E   Q   R   N   A   F   I   Q   S   L   K   D   D   P>
_________o_________o_Z-DOM?N__o_________o_________>

          760         770         780         790         800
aagccaaagc gctaacttgc tagcagaagc taaaaagcta aatgatgctc
ttcggtttcg cgattgaacg atcgtcttcg atttttcgat ttactacgag
  S   Q   S   A   N   L   L   A   E   A   K   K   L   N   D   A>
_________o_________o_Z-DOM?N__o_________o_________>

          810         820         830         840         850
aggcgccgaa aagctcgacc gtagacaaca aattcaacaa agaaacaata
tccgcggctt ttcgagctgg catctgttgt ttaagttgtt tctttgttat
Q   A   P   K>
_________o_>
              S   S   T>
             _________>
                                V   D   N   K   F   N   K   E   T   I>
                               __________AFFI IGA _q_________>

          860         870         880         890         900
caagcgagtc aagagatcag actattacct aacttaaacg gtagacaaaa
gttcgctcag ttctctagtc tgataatgga ttgaatttgc catctgtttt
 Q   A   S   Q   E   I   R   L   L   P   N   L   N   G   R   Q   K>
```

```
_________q_________AFFI IGA _q_________q_________>

       910        920        930        940        950
gcttgccttc atccacagtt tacttgatga cccaagccaa agcgctaact
cgaacggaag taggtgtcaa atgaactact gggttcggtt tcgcgattga
  L  A  F  I  H  S  L  L  D  D  P  S  Q  S  A  N>
_________q_________AFFI IGA _q_________q_________>

                                                  >Xho1
                                                   |
       960        970        980        990       1000
tgctagcaga agctaaaaag ctaaatgatg ctcaggcgcc gaaataactc
acgatcgtct tcgatttttc gatttactac gagtccgcgg ctttattgag
                                                  *>
                                                  ___>
L  L  A  E  A  K  K  L  N  D  A  Q  A  P  K>
_________q_______AFFI IGA ___q_________q___>

gag
ctc
```

[SEQ. ID. NO.: 19] - Recombinant Fiber A7 ZIgG/ZIgG

Sequence Range: 1 to 1003

```
          >EcoR1
            |
            |      10         20         30         40         50
            gaattcgatg aagcgcgcaa gaccgtctga agataccttc aaccccgtgt
            cttaagctac ttcgcgcgtt ctggcagact tctatggaag ttggggcaca
                   M>
                   ___>
```

```
                M   K  R  A   R  P  S  E   D  T  F   N  P  V>
               __j_________j_____TAIL_________j_________>

          60         70         80         90         100
atccatatga cacggaaacc ggtcctccaa ctgtgccttt tcttactcct
taggtatact gtgcctttgg ccaggaggtt gacacggaaa agaatgagga
Y  P  Y  D   T  E  T   G  P  P   T  V  P  F   L  T  P>
_________j_________j__TAIL___j_________j_________>

        110        120        130        140        150
ccctttgtat cccccaatgg gtttcaagag agtccccctg gggtactctc
gggaaacata gggggttacc caaagttctc tcaggggggac cccatgagag
                                                    L  S>
                                                   _____>
 P  F  V   S  P  N  G   F  Q  E   S  P  P   G  V>
_________j_________TAIL_____j_________j____>

        160        170        180        190        200
tttgcgccta tccgaacctc tagttacctc caatggcatg cttgcgctca
aaacgcggat aggcttggag atcaatggag gttaccgtac gaacgcgagt
                                                L  A  L>
                                               _________>
  L  R  L   S  E  P   L  V  T  S   N  G  M>
_________h_____REPEAT 1______h_________>

        210        220        230        240        250
aaatgggcaa cggcctctct ctggacgagg ccggcaacct tacctcccaa
tttacccgtt gccggagaga gacctgctcc ggccgttgga atggagggtt
                                              L   T  S  Q>
                                             __f_________>
K  M  G  N   G  L  S   L  D  E   A  G  N>
_________g_____REPEAT 2_______g_______>

        260        270        280        290        300
aatgtaacca ctgtgagccc acctctcaaa aaaaccaagt caaacataaa
```

```
ttacattggt gacactcggg tggagagttt ttttggttca gtttgtattt
                                                    I  N>
                                                   _____>
 N  V  T   T  V  S  P   P  L  K   K  T  K   S  N>
_________f________REPEAT 3 ____f_________f_____>

      310        320        330        340        350
cctggaaata tctgcacccc tcacagttac ctcagaagcc ctaactgtgg
ggacctttat agacgtgggg agtgtcaatg gagtcttcgg gattgacacc
                                                 L  T  V>
                                                _________>
  L  E  I   S  A  P   L  T  V  T   S  E  A>
_________e_____REPEAT 4_______e_________>

      360        370        380        390        400
ctgccgccgc acctctaatg gtcgcgggca acacactcac catgcaatca
gacggcggcg tggagattac cagcgcccgt tgtgtgagtg gtacgttagt
                                        L  T   M  Q  S>
                                       ___REPEAT 6_____>
A  A  A  A   P  L  M   V  A  G   N  T>
_________d___REPEAT 5_________d_____>

      410        420        430        440        450
caggccccgc taaccgtgca cgactccaaa cttagcattg ccacccaagg
gtccggggcg attggcacgt gctgaggttt gaatcgtaac ggtgggttcc
                                    L  S  I   A  T  Q  G>
                                   ______REPEAT 7_______>
 Q  A  P   L  T  V  H   D  S  K>
__________REPEAT 6__c_________>

      460        470        480        490        500
acccctcaca gtgtcagaag gaaagctagc ccctgacgta gcaagcttac
tggggagtgt cacagtcttc ctttcgatcg gggactgcat cgttcgaatg
                                 L  A>
                                _____a_>
```

```
  P  L  T   V  S  E   G  K>
_________REPEAT 7___b_____>
                                        P  D  V   A  S  L>
                                       ___LUNG SURFACTA____>

       510        520        530        540        550
gacaacaggt agaagccttg caagggcagg tacaacactt acaggcggca
ctgttgtcca tcttcggaac gttcccgtcc atgttgtgaa tgtccgccgt
R  Q  Q  V   E  A  L   Q  G  Q   V  Q  H  L   Q  A  A>
__________LUNG SURFACTANT PROTEIN D TRIM__k__________>

       560        570        580        590        600
tttagccaat acaaaaaggt agagttgttt ccaaacggag ccaagaagct
aaatcggtta tgtttttcca tctcaacaaa ggtttgcctc ggttcttcga
 F  S  Q   Y  K  K  V   E  L  F   P  N  G>
_____LUNG SURFACTANT PROTEIN D TRIM_______>
                                              A  K  K  L>
                                            _l_________>

                                    >ClaI
                                      |
       610        620        630    |    640        650
gaacgacgcc caggccccca agagcgaccc atcgatcgta gacaacaaat
cttgctgcgg gtccgggggt tctcgctggg tagctagcat ctgttgttta
  N  D  A   Q  A  P   K  S  D>
__________SPA 48-60__l________>
                                       P   S  I>
                                     __m_______>
                                                V   D  N  K>
                                              ___o__________>

       660        670        680        690        700
tcaacaaaga acaacaaaac gcgttctatg agatcttaca tttacctaac
agttgtttct tgttgttttg cgcaagatac tctagaatgt aaatggattg
F  N  K  E   Q  Q  N   A  F  Y   E  I  L  H   L  P  N>
```

```
          _____o_____o_Z-DOM?N__o_____o_____>

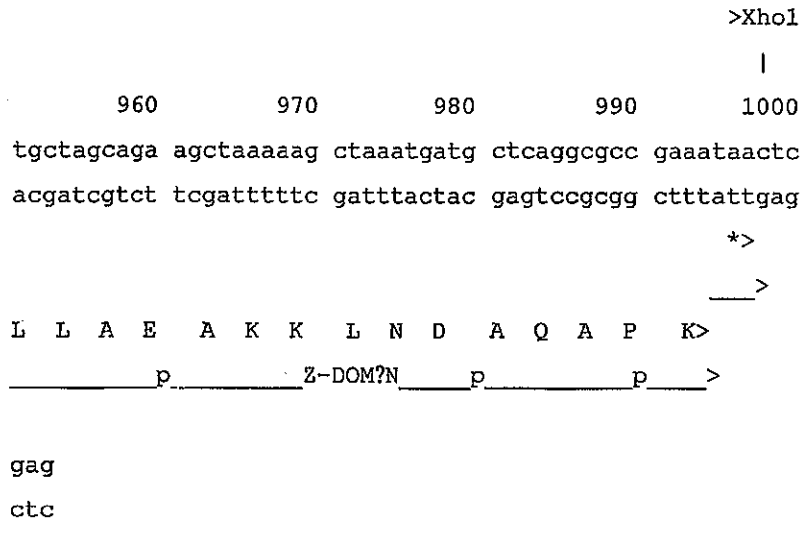
                710        720        730        740        750
          ttaaacgaag aacaacgaaa cgccttcatc caaagtttaa aagatgaccc
          aatttgcttc ttgttgcttt gcggaagtag gtttcaaatt ttctactggg
           L  N  E   E  Q  R  N   A  F  I   Q  S  L   K  D  D  P>
          _____o_____o_Z-DOM?N__o_____o_____>

                760        770        780        790        800
          aagccaaagc gctaacttgc tagcagaagc taaaaagcta aatgatgctc
          ttcggtttcg cgattgaacg atcgtcttcg atttttcgat ttactacgag
           S  Q  S   A  N  L   L  A  E  A   K  K  L   N  D  A>
          _____o_____o_Z-DOM?N__o_____o_____>

                810        820        830        840        850
          aggcgccgaa aagctcgacc gtagacaaca aattcaacaa agaacaacaa
          tccgcggctt ttcgagctgg catctgttgt ttaagttgtt tcttgttgtt
          Q  A  P  K>
          _____o_>
                                V  D  N   K  F  N  K   E  Q  Q>
                               _____p_Z-DOM?N__p_____>
                          S  S  T>
                         _____>

                860        870        880        890        900
          aacgcgttct atgagatctt acatttacct aacttaaacg aagaacaacg
          ttgcgcaaga tactctagaa tgtaaatgga ttgaatttgc ttcttgttgc
           N  A  F   Y  E  I  L   H  L  P   N  L  N   E  E  Q  R>
          _____p_____p_Z-DOM?N__p_____p_____>

                910        920        930        940        950
          aaacgccttc atccaaagtt taaaagatga cccaagccaa agcgctaact
          tttgcggaag taggtttcaa attttctact gggttcggtt tcgcgattga
           N  A  F   I  Q  S   L  K  D  D   P  S  Q   S  A  N>
          _____p_____p_Z-DOM?N__p_____p_____>
```



SEQ I.D. No.: 47 - Anti β-galactosidase single chain Fv fragment

ATGGCCGAGGTGCAGCTGGTGGAGTCTGGGGGAAGCCTGGTCAAGCCTGGGGGG
TCCCTGAGACTCTCCTGTGCAGCCTCTGGATTCACCTTCAGTAACTATAGCATGAAC
TGGGTCCGCCAGGCTCCAGGGAAGGGGCTGGAGTGGATCTCATCCATTAGTGGTAG
TAGTAGATACATATACTACGCAGACTTCGTGAAGGGCCGATTCACCATCTCCAGAGAC
AACGCCACGAACTCACTGTACCTGCAAATGAACAGCCTGAGAGCCGAGGACACGGCTG
TTTATTACTGTGTGAGATCCAGTATTACGATTTTTGGTGGCGGTATGGACGTCTGG
GGCAGAGGCACCCTGGTCACCGTCTCCTCAGGTGGAGGCGGTTCAGGCGGAGGTGG
CAGCGGCGGTGGCGGATCGCAGTCTGTGCTGACTCAGCCTGCCTCCGTGTCTGGGT
CTCCTGGACAGTCGATCACCATCTCCTGCGCTGGAACCAGCAGTGACGTTGGTGGT
TATAACTATGTCTCCTGGTACCAACAACACCCAGGCAAAGCCCCCAAACTCATGAT
TTATGAGGACAGTAAGCGGCCCTCAGGGGTTTCTAATCGCTTCTCTGGCTCCAAGT
CTGGCAACACGGCCTCCCTGACAATCTCTGGGCTCCAGGCTGAGGACGAGGCTGAT
TATTACTGCAGCTCATATACAACCAGGAGCACTCGAGTTTTCGGCGGAGGGACCAA
GCTGGCCGTCCTAGGTGCGGCCGCAGAACAAAAACTCATCTCAGAAGAGGATCTGA
ATGGGGCCGCACATCACCATCATCACCAT

Claims:

1. A modified virus comprising one or more non-native polypeptides, which polypeptide comprises one or more framework moieties which possesses a folded three dimensional structure, each framework moiety containing one or more binding moieties, which polypeptide is capable of being expressed in the cytoplasm and nucleus of a mammalian host cell in a conformation which is maintained in the absence of a ligand for said binding moieties, said conformation allowing said binding moieties subsequently to bind with said ligand, and which polypeptide is capable of transport though the nuclear membrane, wherein said modified virus has an altered tropism conferred by said binding moieties.

2. The modified virus of claim 1 which is derived from a virus selected from the group consisting of adenoviruses, retroviruses, lentiviruses, adeno-associated viruses, Reoviridae, Picornaviridae, Parvoviridae, Papovaviridae and Caliciviridae.

3. The modified virus of claim 1 or claim 2 which is derived from human adenovirus

4. The modified virus of any one of claims 1 to 3 which is derived from human adenovirus serotype 5.

5. The modified virus of any one of claims 1 to 4 wherein said non-native polypeptide replaces, is incorporated into, or forms a fusion protein with, a viral protein component of the wild type virus.

6. The modified virus of any one of claims 1 to 5 wherein said non-native polypeptide and/or the viral component protein comprising the non-native polypeptide is soluble in a cellular environment.

7. The modified virus of claim 6 wherein greater than 30% of the non-native polypeptide or the viral component protein comprising the non-native polypeptide is present in the soluble fraction of cell lysates of cells expressing the non-native polypeptide or viral component protein comprising the non-native polypeptide.

8. The modified virus of any one of claims 5 to 7 wherein said viral protein component is an adenoviral fiber protein.

9. The modified virus of claim 8 wherein said non-native polypeptide is incorporated into an adenoviral fiber protein such that the wild-type fiber knob or cell binding domain thereof is removed.

10. The modified virus of any one of claims 1 to 9 wherein said non-native polypeptide comprises one or more further elements that mimic the native structure or function of a viral component protein in which said non-native polypeptide is incorporated or which said non-native polypeptide replaces.

11. The modified virus of any one of claims 1 to 10 wherein said non-native polypeptide does not contain any di-sulphide bonds.

12. The modified virus of any one of claims 1 to 11 wherein the non-native polypeptide comprises one or more α-helical structures.

13. The modified virus of any one of claims 1 to 12 wherein said non-native polypeptide comprises a framework moiety derived from an antibody which is capable of productive folding in the cytoplasm and subsequent

transport into the cell nucleus.

14. The modified virus of claim 13 wherein said framework moiety comprises one or more binding moieties derived from a CDR of a further antibody.

15. The modified virus of any one of claims 1 to 12 wherein said non-native polypeptide is or comprises a combinatorial protein or an affibody.

16. The modified virus of any one of claims 1 to 15 wherein said non-native polypeptide comprises one or more framework moieties derived from a bacterial receptor.

17. The modified virus of any one of claims 1 to 12 or 16 wherein said non-native polypeptide comprises one or more binding moieties which are present within one or more loops of a helical bundle and/or one or more of the loops connecting these bundles.

18. The modified virus of claim 16 wherein said non-native polypeptide comprises one or more framework moieties derived from the immunoglobulin binding Z-domain from staphylococcal protein A or an immunoglobulin binding domain from streptococcal protein G.

19. The modified virus of any one of claims 1 to 18 wherein said non-native polypeptide comprises one or more framework moieties which comprise a non native trimerisation motif.

20. The modified virus of claim 19 wherein said non-native polypeptide comprises one or more framework moieties which comprise the neck region peptide from human lung surfactant protein D.

21. The modified virus of any one of claims 1 to 20 which comprises two or more different non-native polypeptides.

22. The modified virus of claim 21 which comprises a first non-native polypeptide which binds a target cell and a second non-native polypeptide which binds a production cell or permissive cell.

23. The modified virus of any one of claims 1 to 22 wherein said non-native polypeptide comprises a cleavage site positioned in a location that enables a binding moiety of the non-native polypeptide to be cleaved from the modified virus.

24. The modified virus of any one of claims 1 to 23 wherein said virus comprises a modified viral component comprising said non-native polypeptide and a corresponding unmodified viral component e.g. a wild-type fiber and a modified fiber.

25. The modified virus of any one of claims 1 to 24 wherein said non-native polypeptide comprises a binding moiety capable of binding to a cell specific ligand which may optionally be Prostate Specific Membrane Antigen, EGF receptor, Her-2/Neu, VEGF receptor, CD22, gp120, MHC/peptide complexes or membrane structures or surface molecules expressed or present on proliferating cells, tumor cells or virus infected cells.

26. The modified virus of any one of claims 1 to 25 which further comprises a site for insertion of one or more desired therapeutic genes or nucleic acid molecules.

27. The modified virus of claim 26 which comprises

transgenes encoding cytosine deaminase and/or uracil phosphoribosyl transferase either as separate genes or together as a bifunctional fusion gene.

28. The modified virus of any one of claims 1 to 27 which further comprises a viral component which is replaced with an equivalent component from a different serotype or is modified such that binding of said virus by antibodies pre-formed to the wild type component is reduced.

29. The modified virus of claim 28 wherein said viral component is a hexon protein.

30. A modified viral protein comprising a non-native polypeptide as defined in any one of claims 1 to 25.

31. A cell containing a modified virus or viral protein as defined in any of the claims 1 to 30.

32. A permissive cell for a modified virus as defined in any one of claims 1 to 29 which is capable of being cultured to propagate said modified virus.

33. A method for producing a modified virus as defined in any one of claims 1 to 29 in cell culture, comprising the steps of:

i) genetically modifying a virus to produce a modified virus comprising one or more non-native polypeptides, which polypeptide comprises one or more framework moieties each containing one or more binding moieties, which polypeptide is capable of being expressed in the cytoplasm and nucleus of a mammalian host cell in a conformation which is maintained in the absence of a ligand for said binding moieties, said conformation allowing said binding moieties subsequently to bind with said ligand, and which

polypeptide is capable of transport though the nuclear membrane, wherein said modified virus has an altered tropism conferred by said binding moieties;

ii) infecting permissive cells with said modified virus;

iii) culturing said cells to produce the virus;

and

iv) harvesting, and optionally, purifying the modified virus produced.

34. A method of regulating the replication of a modified virus as defined in any one of claims 1 to 29 comprising the steps of:

i) constructing a modified virus such that a cleavage site is positioned between a binding moiety required for cell infection and the remainder of the recombinant viral component of which the binding moiety forms part; and,

ii) bringing said recombinant virus into contact with a cleavage agent or cleavage means capable of cleaving said binding moiety from said viral component and thereby preventing the recombinant virus from undergoing further infection cycles.

35. A method of determining the suitability of a non-native polypeptide for use in the preparation of a viral vector by determining its solubility in a cell system.

36. A method as claimed in claim 35 which comprises the steps of:

i) expressing said non-native polypeptide or a viral component protein comprising said non-native polypeptide in permissive cells;

ii) subjecting the cells to lysis to produce a cell lysate;

iii) separating the soluble and insoluble fractions of the cell lysate;

iv) analysing the soluble and insoluble fractions of the cell lysate for the content of said non-native polypeptide or viral component protein comprising said non-native polypeptide;
and,
v) comparing the relative content of said non-native polypeptide or viral component protein comprising said non-native polypeptide in the soluble and insoluble fractions

37. A method of assaying the solubility of a non-native polypeptide as defined in any one of claims 1 to 25 or a modified viral protein as defined in claim 30, which method comprises the steps of:
i) expressing said non-native polypeptide or a viral component protein comprising said non-native polypeptide in permissive cells;
ii) subjecting the cells to lysis to produce a cell lysate;
iii) separating the soluble and insoluble fractions of the cell lysate;
iv) analysing the soluble and insoluble fractions of the cell lysate for the content of said non-native polypeptide or viral component protein comprising said non-native polypeptide;
and,
v) comparing the relative content of said non-native polypeptide or viral component protein comprising said non-native polypeptide in the soluble and insoluble fractions.

38. The modified virus of any one of claims 1 to 29 for use in therapy.

39. Use of the modified virus of any one of claims 1 to

29 in the preparation of a medicament for the treatment of tumour cells or proliferating cells.

40. A pharmaceutical composition comprising a modified virus of any one of claims 1 to 29 and a pharmaceutically acceptable carrier or excipient.