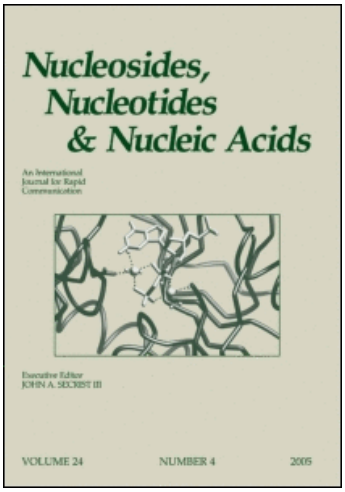

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Cloning and Expression of Collagen-Derived Protein in E. Coli And Baculovirus System

S. Schlosser[a]; G. Gröger[a]; O. Holfeld[a]; H. Seliger[a]
[a] Sektion Polymere, Universität Ulm, Ulm, Germany

To cite this Article Schlosser, S. , Gröger, G. , Holfeld, O. and Seliger, H.(1998) 'Cloning and Expression of Collagen-Derived Protein in E. Coli And Baculovirus System', Nucleosides, Nucleotides and Nucleic Acids, 17: 9, 1515 — 1522
To link to this Article: DOI: 10.1080/07328319808004681
URL: http://dx.doi.org/10.1080/07328319808004681

NUCLEOSIDES & NUCLEOTIDES, 17(9-11), 1515-1522 (1998)

CLONING AND EXPRESSION OF COLLAGEN-DERIVED PROTEIN IN E.COLI AND BACULOVIRUS SYSTEM

S. Schlosser, G. Gröger, O. Holfeld, and H. Seliger*
Sektion Polymere, Universität Ulm
Albert Einstein-Allee 11, D-89069 Ulm, Germany

ABSTRACT: Using appropriately designed primers we amplified a 783 bp fragment of the triple helical portion of human α1-collagen(I) by PCR methods. The expression of this biosynthetic collagen gene and a modified one containing cell adhesion sequences was carried out in a prokaryotic (*E.coli*) and a eukaryotic (*baculovirus*) system. The expression products were purified by affinity chromatography.

Materials on the basis of cell matrix proteins such as collagen, laminin or fibronectin are of great interest for potential medical applications [1]. The supply of many eukaryotic proteins, such as human collagen, is often limited by their low availability from biological sources. Gene cloning and expression in E.coli can provide an abundant source of these proteins [2]. On the other hand, in the last years heterologous gene expression in insect cells using baculovirus vectors [3] is also acknowledged as a powerful system for recombinant protein production. The ability of insect cells to perform most of the sophisticated post-translational modifications found in higher eukaryotic cells makes the system an interesting tool for the production of a wide variety of proteins [4].

Collagen is the most abundant protein in human body. Various types of collagen are responsible for the functional integrity of tissues as bone, cartilage and skin. All collagens have in common a triple helical domain, in which each of the three α-helices is characterized by the Gly-X-Y repeating sequence. The X and Y positions are frequently proline and hydroxyproline respectively [5].

Using this structural information about collagens we are studying the cloning and expression of recombinant biosynthetic human collagens in prokaryotic (E.coli) and eukaryotic (insect cells) hosts[6,7].

EXPERIMENTAL

Preparation and cloning of the biosynthetic gene(s):

Coll: A fragment of the gene encoding the crystalline part of human α1-collagen (type I) was amplified by PCR using a cDNA library of human placenta, the sense primer CAG CCA TGG CCC TCT GGT CCT CGT GGT CT and the antisense primer T GCA CTG CAG TCC TCG CTT TCC TTC CTC TC. The primers were synthesized in a DNA Synthesizer (Pharmacia Gene Assembler Plus) according to the phosphoramidite method[8]. The gene construct we obtained (*Coll*; FIG. 1, sequence B) was first cloned into the vector pGEM-5Zf(+)T (Promega) and then into pTrc99A (Pharmacia) using the restriction sites for NcoI and PstI. For expression in E.coli *Coll* was subcloned into the vector pET29b(+) (Novagen) using the restriction sites for NcoI and HindIII. For expression in the Sf9 insect cells *Coll* was subcloned in the baculovirus vector pBlueBacHis2B (Invitrogen).

RGDS-Coll: In another approach we modified the biosynthetic collagen gene (*Coll*) by adding an oligonucleotide sequence coding for the amino acid sequence $(RGDS)_4$ at its N-terminus in order to enhance cell adhesion to Coll-protein [9-11]. Restriction of *Coll* in pBlueBacHis2B with KpnI and insertion of the oligonucleotide sequence coding for the $(RGDS)_4$ amino acid sequence led to the gene construct *RGDS-Coll* (FIG.1, sequence A+B).

Sequence analysis was performed with amplification primers using dideoxy cycle sequencing and dye terminators in an automated sequencer (Perkin-Elmer, Applied Biosystems)(FIG.1).

Expression in E.coli:

For prokaryotic expression *Coll* in pET-29b(+) was transformed in E.coli strains HMS174(DE3) and BLR(DE3)pLysS by electroporation (E.coli Pulser™, Biorad). This vector contains a His•Tag (recognition site for Ni^{2+}-NTA conjugate), which allows

purification of the expression product via Ni^{2+}-NTA agarose affinity chromatography and detection of the recombinant protein by Western blotting methods.

A single colony was inoculated into 3-4 ml Lennox L Broth (LB) medium containing the appropriate antibiotic (see TABLE 1) and incubated overnight at 37°C with shaking.

Ca. 100 µl of the starter culture were inoculated into 50 ml LB medium and grown until the density of the cell suspension reached an OD_{600} of 0.6-0.7 (ca. 4 hours). Expression of the target protein was induced by addition of 1-2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and incubation for another 2 hours.

Expression in Sf9 insect cells:

For gene amplification the genes for *Coll* and *RGDS-Coll* inserted into pBlueBacHis2B were first transformed in E.coli strain DH5α by electroporation (E.coli Pulser™, Biorad). Each of the amplified genes was then recombined with the Autographa californica Multicapsid Nuclear Polyhedrosis Virus (AcMNPV, Invitrogen) used for transfection of the Sf9 insect cells. The transfection of these cells and the expression of the biosynthetic genes was performed according to Ref. 4 and the instruction manual of Invitrogen Corp. (see TABLE 2).

Purification:

The cell culture was harvested by centrifugation (10 min, 5000 x *g*, 4°C). The cell pellet was resuspended in 1/10 cell culture volume cell-lysis buffer (10 mM Tris HCl, pH 8.0, 1% Triton·X-100) and incubated on ice for 30 min. During incubation the cell suspension was sonicated 3-4 times for 20-30 sec in a sonication bath. To reduce the viscosity the cell lysate was passed 3 times through a 0.4mm syringe. The insoluble fraction was removed by centrifugation (20 min, 10 000 x *g*, 4°C).

Via the C-terminal His•Tag sequence both of the biosynthetic genes bind to divalent cations (*e.g.* Ni^{2+}) immobilized on Ni^{2+}-NTA agarose (Qiagen). 150 µl slurry of Ni^{2+}-NTA agarose was added to the soluble fraction of the cell lysate, and the suspension was incubated for ca. 1 hour at room temperature with shaking. The suspension was then transferred to a column, and the resin was washed with 20 ml of buffer A (10 mM Tris, 300 mM NaCl, 0.25%Triton·X-100, pH 8.0) and 20ml of 20mM imidazole in buffer B (10 mM Tris, 300 mM NaCl, 0.25%Triton·X-100; pH 6.3). After unbound proteins had been

TABLE 1: Expression of *Coll* in two different E.coli strains.

strain	*BLR(DE3)pLysS*	*HMS174(DE3)*
antibiotic resistance	chloramphenicol (34 µg/ml) tetracycline (12.5 µg/ml)	none
expression vector	pET 29b(+) incl.(His•Tag)	pET 29b(+) incl.(His•Tag)
selection	kanamycin (30 µg/ml)	-
promoter	T7lac	T7lac
N-terminal leader	S•Tag	S•Tag
induction of expression	by IPTG (1-2 mM)	by IPTG (1-2 mM)

TABLE 2: Cloning, transfection and expression of *Coll* and *RGDS-Coll* in Sf9 cells.

	Cloning	*Transfection*	*Expression*
strain	E.coli, DH5a	insect cells, Sf9	insect cells, Sf9
vector *virus*	pBlueBacHis2B (His•Tag) ---	--- recomb. AcMNPV	--- recomb. AcMNPV
selection	ampicillin (50 µg/ml)	lethal deletion by recombination	---
promoter	T7lac		polyhedrin

washed away, the target protein fraction was recovered from the column by elution with 300µl of 250mM imidazole in buffer B.

RESULTS AND DISCUSSION

A fragment of a cDNA library of human placenta collagen was amplified by PCR methods using the synthetic primers described above. We obtained a gene construct consisting of 783 bp (*Coll*; FIG. 1, sequence B).

Adding the oligonucleotide (see FIG. 1, sequence A) coding for the amino acid sequence $(RGDS)_4$ we obtained the modified biosynthetic collagen gene *RGDS-Coll* (FIG.1, sequence A+B).

These gene constructs were cloned and expressed in E.coli and in insect cells (using a baculovirus vector). The expression products were isolated after cell lysis and purified via Ni^{2+}-NTA affinity chromatography.

A

KpnI
GGT ACC CGT GGT
CCA TGG GCA CCA
Gly Thr Arg Gly
R G

GAC TCT CGT GGT GAC TCT CGT GGT GAC TCT CGT GGT GAC TCT CGT GTT GAC TCT GGT
CTG AGA GCA CCA CTG AGA GCA CCA CTG AGA GCA CCA CTG AGA GCC CAA CTG AGA CCA
Asp Ser Arg Gly Asp Ser Arg Gly Asp Ser Arg Gly Asp Ser Arg Val Asp Ser Gly
D S R G D S R G D S R G D S

B

NcoI
ACC ATG GGT CCC CGA GGT CCC CCA GGT CCC CCT GGA AAG AAT GGA GAT GAT GGG GAA GCT
TGG TAC CCA GGG GCT CCA GGG GGT CCA GGG GGA CCT TTC TTA CCT CTA CTA CCC CTT CGA
Thr Met Gly Pro Arg Gly Pro Pro Gly Pro Pro Gly Lys Asn Gly Asp Asp Gly Glu Ala

GGA AAA CCT GGT CGT CCT GGT GAG CGT GGG CCT CCT GGG CCT CAG GGT GCT CGA GGA
CCT TTT GGA CCA GCA GGA CCA CTC GCA CCC GGA GGA CCC GGA GTC CCA CGA GCT CCT
Gly Lys Pro Gly Arg Pro Gly Glu Arg Gly Pro Pro Gly Pro Gln Gly Ala Arg Gly

TTG CCC GGA ACA GCT GGC CTC CCT GGA ATG AAG GGA CAC AGA GGT TTC AGT GGT TTG
AAC GGG CCT TGT CGA CCC GAG GGA CCT TAC TTC CCT GTG TCT CCA AAG TCA CCA AAC
Leu Pro Gly Thr Ala Gly Leu Pro Gly Met Lys Gly His Arg Gly Phe Ser Gly Leu

GAT GGT GCC AAG GGA GAT GCT GGT CCT GCT GGT CCT AAG GGT GAG CCT GGC AGC CCT
CTA CCA CGG TTC CCT CTA CGA CCA GGA CGA CCA GGA TTC CCA CTC GGA CCG TCG GGA
Asp Gly Ala Lys Gly Asp Ala Gly Pro Ala Gly Pro Lys Gly Glu Pro Gly Ser Pro

GGT GAA AAT GGA GCT CCT GGT CAG ATG GGC CCC CGT GGC CTG CCT GGT GAG AGA GGT
CCA CTT TTA CCT CGA GGA CCA GTC TAC CCG GGG GCA CCG GAC GGA CCA CTC TCT CCA
Gly Glu Asn Gly Ala Pro Gly Gln Met Gly Pro Arg Gly Leu Pro Gly Glu Arg Gly

CGC CCT GGA GCC CCT GGC CCT GCT GGT GCT CGT GGA AAT GAT GGT GCT ACT GGT GCT
GCG GGA CCT CGG GGA CCG GGA CGA CCA CGA GCA CCT TTA CTA CCA CGA TGA CCA CGA
Arg Pro Gly Ala Pro Gly Pro Ala Gly Ala Arg Gly Asn Asp Gly Ala Thr Gly Ala

GCC GGG CCC CCT GGT CCC ACC GGC CCC GCT GGT CCT CCT GGC TTC CCT GGT GCT GTT
CGG CCC GGG GGA CCA GGG TGG CCG GGG CGA CCA GGA GGA CCG AAG GGA CCA CGA CAA
Ala Gly Pro Pro Gly Pro Thr Gly Pro Ala Gly Pro Pro Gly Phe Pro Gly Ala Val

GGT GCT AAG GGT GAA GCT GGT CCC CAA GGG CCC CGA GGA TCT GAA GGT CCC CAG GGT
CCA CGA TTC CCA CTT CGA CCA GGG GTT CCC GGG GCT CCT AGA CTT CCA GGG GTC CCA
Gly Ala Lys Gly Glu Ala Gly Pro Gln Gly Pro Arg Gly Ser Glu Gly Pro Gln Gly

GTC CTG GGT GAG CCT GGT CCC CCT GGC CCT GCT GGT GCT GCT GGC CCT GCT GGA AAC
CAG GAC CCA CTC GGA CCA GGG GGA CCG GGA CGA CCA CGA CGA CCG GGA CGA CCT TTG
Val Leu Gly Glu Pro Gly Pro Pro Gly Pro Ala Gly Ala Ala Gly Pro Ala Gly Asn

CCT GGT GCT GAT GGA CAG CCT GGT GCT AAA GGT GCC AAT GGT GCT CCT GGT ATT GCT
GGA CCA CGA CTA CCT GTC GGA CCA CGA TTT CCA CGG TTA CCA CGA GGA CCA TAA CGA
Pro Gly Ala Asp Gly Gln Pro Gly Ala Lys Gly Ala Asn Gly Ala Pro Gly Ile Ala

(continued)

FIG. 1: Sequence of the biosynthetic *Coll* and the modified *RGDS-Coll* gene and corresponding amino acid sequence. *RGDS-Coll:* Sequence A directly followed by B; *Coll:* Sequence B. The restriction sites for KpnI, NcoI, PstI and HindIII used for cloning are also indicated.

```
GGT GCT CCA GGT TTC CCT GGT GCC CGA GGC CCC TCT GGA CCC CAA GGC CCA GGC GGC
CCA CGA GGT CCA AAG GGA CCA CGG GCT CCG GGG AGA CCT GGG GTT CCG GGT CCG CCG
Gly Ala Pro Gly Phe Pro Gly Ala Arg Gly Pro Ser Gly Pro Gln Gly Pro Gly Gly

CCT CCT GGT CCC AAG GGT AAC AGC GGT GAA CCT GGT GCT CCT GGC AGC AAA GGA GAC
GGA GGA CCA GGG TTC CCA TTG TCG CCA CTT GGA CCA CGA GGA CCG TCG TTT CCT CTG
Pro Pro Gly Pro Lys Gly Asn Ser Gly Glu Pro Gly Ala Pro Gly Ser Lys Gly Asp

ACT GGT GCT AAG GGA GAG CCT GGC CCT GTT GGT GTT CAA GGA CCC CCT GGC CCT GCT
TGA CCA CGA TTC CCT CTC GGA CCG GGA CAA CCA CAA GTT CCT GGG GGA CCG GGA CGA
Thr Gly Ala Lys Gly Glu Pro Gly Pro Val Gly Val Gln Gly Pro Pro Gly Leu Ala

                                PstI                HindIII
GGA GAG GAA GGA AAG CGA GGA CTG CAG GCA TGC AAG CTT
CCT CTC CTT CCT TTC GCT CCT GAC GTC CGT ACG TTC GAA
Gly Glu Glu Gly Lys Arg Gly Leu Gln Ala Cys Lys Ser
```

FIG. 1 Continued

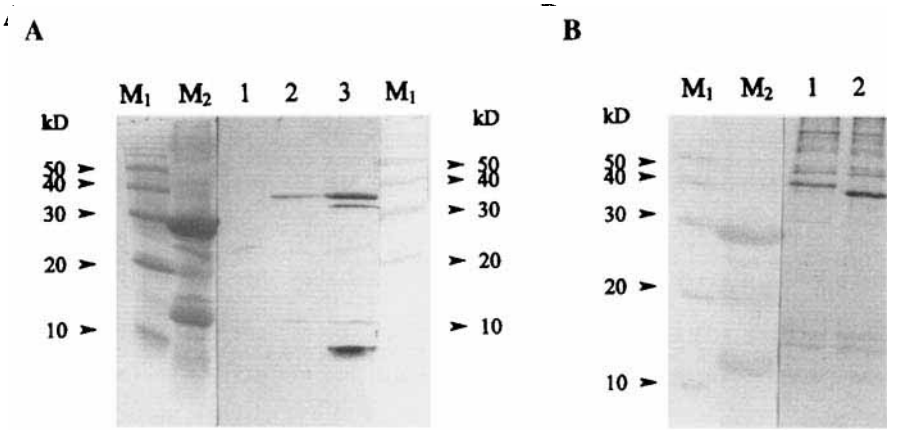

FIG. 2: Western blot of *Coll* in E. coli (A) and *Coll* and *RGDS-Coll* in Sf9 insect cells (B) after purification via Ni^{2+} NTA agarose affinity chromatography.

A

M_1: molecular weight marker;10 kD ladder; stain with amido black
M_2: Rainbow marker; stain with amido black
1: uninduced control (elution fraction)
2: *Coll* (HMS174(DE3); elution fraction)
3: *Coll* (BLR(DE3)pLysS; elution fraction)
M'_1: molecular weight marker; 10kD ladder; stain with Ponceau S

B

M_1: molecular weight marker; 10 kD ladder; stain with Ponceau S
M_2: Rainbow marker; stain with Ponceau S
1: *RGDS-Coll*; elution fraction
2: *Coll* ; elution fraction

The eluted fraction(s) were analyzed by electrophoresis on a 16% SDS-polyacrylamide-gel. For identification the purified expression product(s) were transferred onto nitrocellulose and incubated with Ni^{2+}-NTA (nitrilotriacetic acid) conjugate (Qiagen) which binds to the C-terminal His•Tag. The target protein(s) were visualized by BCIP (5-bromo-4-chloro-3-indolyl phosphate) and NBT (nitro blue tetrazolium) (Sigma). Visualization of the molecular weight standards (Rainbow marker, Amersham; 10 kD ladder, GIBCO BRL) occurred directly after the transfer onto nitrocellulose by staining with Ponceau S (Sigma) or amido black (Sigma) (FIG. 2 A,B).

The *Coll* protein, expressed in E.coli (FIG.2A) and in Sf9 cells (FIG.2B) appears as a bold band with an apparent molecular weight of ca. 35kD, which corresponds to the one calculated from the amino acid sequence (FIG. 1, sequence B).

The expressed *RGDS-Coll* in Sf9 cells has an apparent molecular weight of ca. 40kD, which is expected because of the additional $(RGDS)_4$ sequence at the N-terminus of the (*Coll*) protein (s. FIG. 1, sequence A+B).

In the Western blots the elution fractions (FIG. 2A, lanes 1-3 and 2B, lanes 1-2) contain some protein contamination, which is probably due to the weak binding of other histidine-containing proteins to the resin and to the Ni^{2+}-NTA conjugate. The protein bands appearing at ca. 8kD and ca. 31kD (FIG. 2A, lane 3) could be a hint that in E.coli BLR(DE3)pLysS some portion of the expression product is degraded by proteolysis or that the expression of *Coll* is partially incomplete in this expression system. Therefore, a second purification step following the Ni^{2+}-affinity chromatography, e.g. gel-filtration chromatography, seems to be necessary.

Further investigations are in progress to analyze the biophysical properties of these recombinant biosynthetic proteins i.e. their ability to adopt a triple helical conformation, their degree of hydroxylation (especially of the genes expressed in the eukaryotic system), their biocompatibility etc.

ACKNOWLEDGMENT

Financial support by the Deutsche Forschungsgemeinschaft (DFG) in the frame of the Forschergruppe „Biomechanik und Biologie der operativen Behandlung der Verletzungen des Bewegungsapparates“ is gratefully acknowledged.

REFERENCES

1. Cappello, J.; Grissman, J.W. *Polymer Preprints*, **1990**, *30*, 193-194.
2. Marston, F.A.O., Biochem. J., **1986**, *240*, 1-12.
3. O'Reilly, D.R.; Miller, L.K.; Luckow, V.A., Baculovirus Expression Vectors, **1992,** W.H. Freeman, New York.
4. Tomita, M.; Ohkura, N.; Ito, M.; Kato, T.; Royce, P.M.; Kitajima, T., *Biochem. J.*, **1995**, *312*, 847-853.
5. Kadler, K., *Protein Profiles*, **1995**, *Vol. 2, Issue 5*, Academic Press, ISSN 1070-3667.
6. For preliminary communications see e.g.: Földes-Papp, Z.; Seliger, H.; Ramalho-Ortigao, J.F.; *Abstracts, VIth Colloquium on Biomaterials, Aachen*, **1992;** Ramalho-Ortigao, J.F.; Földes-Papp, Z.; Brüninghoff, A.; Seliger, H.; *Abstracts, The John Ugelstad Conference II, Cambridge, Mass.*, **1993,** 35.
7. Brüninghoff, A.; Schlosser, S.; Gröger, G.; Seliger, H. *Nucleosides & Nucleotides,* **1997**, *16*, 875-882.
8. Beaucage, S.L.; Caruthers, M.H., *Tetrahedron Lett.*, **1981**, *22*, 1859-1862.
9. Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A., *J.Biol.Chem.*, **1982**, *257*, 9593-9597.
10. Pierschbacher, M.D.; Ruoslahti, E., *Nature,* **1988**, *309*, 30-33.
11. Hynes, R.O., *Cell,* **1992**, *69*, 11-25.