Scientific paper

Influence of Primary Gene Structure on Secretion of Recombinant TNFα From *Aspergillus niger*

Nada Kraševec* and Radovan Komel

National Institute of Chemistry, POB 660, SI-1001 Ljubljana, Slovenia

** Corresponding author: E-mail: nada.krasevec@ki.si; fax: (+386)-1-4760-300*

Received: 12-11-2007

Abstract

We studied secretion of human cytokine tumour necrosis factor alpha (TNFα) by the filamentous fungus *Aspergillus niger*. When the gene for glucoamylase was fused with TNFα, optimized for expression in *E. coli* (TNFE), TNFα was detected only inside cells, in the form of an uncleaved fusion protein. Non-optimized codon usage of TNFE proved negative influence on the expression of specific mRNA. The influence of the cluster of unfavourable codons was studied further by site-directed mutagenesis in which these codons were optimized by eukaryotic codon usage, as present in human TNFα (TNFH) or in fungal glucoamylase, however, no improvement was observed. Processed TNFα was secreted only when entire TNFH replaced TNFE. If the codons of the cluster within TNFH were shifted back to bacterial codon usage, the secretion of TNFα dropped down. Due to the great impact of codon usage observed in our experiments, human cDNA seems appropriate for expression in *A. niger*, harbouring codons complementary enough to the fungal host specific codon usage. Methods of secondary structure prediction for mRNA proved being a helpful tool in designing gene constructs including those containing heterologous intron(s).

Keywords: Filamentous fungi; heterologous protein secretion; human tumour necrosis factor α; kexin processing; codon usage; intron.

1. Introduction

During the past few decades tumour necrosis factor α (TNFα) has been extensively studied due to its important physiological role. It is also a therapeutic anticancer drug. Systemic application, however, has always been associated with complications during treatment. Encouraging results were achieved by applying a limb perfusion technique,[1] and the drug Beromun®, containing tasonermin (TNFα–1a) as an active ingredient, was registered. The TNFα monomer is a 17 kDa non-glycosylated protein, containing a single intra-molecular disulfide bond, but it exists in solution as a compact, bell-shaped homotrimer, which is considered to be the biologically active form of this important cytokine. However, in spite of successful expression of TNFα in several different host organisms, its price has remained high. Filamentous fungi present a good choice for heterologous expression, because of their ability to secrete proteins in large quantities, and other benefits compared to other heterologous expression systems.[2] Although several successful strategies for secreting heterologous proteins have been developed and several milligrams of a heterologous protein per litre could be obtained, the gap between production of homologous and heterologous proteins still persists. Strategies to decrease this difference in protein production should be searched for at transcriptional, as well as (post)translational levels.[3–9]

The proteins of interest to be produced by filamentous fungi fall roughly into two categories: industrial enzymes and gene products from higher organisms, which are often of pharmaceutical value. To date, there are several examples showing that gene products originating from other fungi can be successfully produced up to the gram per litre level.[10] The human proteins proinsulin, lymphotoxin α, tissue plasminogen activator (t-PA), interleukin 6 and humanized antibodies were expressed in the fungus *Aspergillus niger*.[11–15]

The influence of codon usage on heterologous protein expression has not been studied systematically in fungi. In the case of the plant *Cymamopsis tetragonoloba* α-galactosidase, a small amount of full size mRNA was de-

tected in fungus *A. awamori* only after expressing the gene with codon usage optimized for yeast.[16] Van Gemeren et al. reported the influence of codon usage on secretion of fungal cutinase analogues.[17] Secretion of the plant *Thaumatoccocus daniellii* sweet-tasting protein thaumatin by recombinant strains of *A. awamori* was achieved after a synthetic gene optimized for filamentous fungi was used.[18,19] For expressing GFP in several fungi, the synthetic S65T-GFP optimized for expression in plants was the best choice.[20,21] The highest production of insulin in *A. niger* was recorded after adapting 30 codons of a synthetic gene for proinsulin to those most commonly used in glucoamylase G1.[11] Unsuccessful transcription of the xylanase gene *xynB* from the thermophilic bacterium *Dictyoglomus thermophilum* during expression in the fungus *Trichoderma reesei* was overcome by using a synthetic gene, in which 20 codons were adapted to codon usage of highly expressed genes of *T. reesei*.[22] Expression of a synthetic copy of the bovine chymosin gene led to more efficient production of chymosin by the fungus *A. awamori*.[23] In case of humanized antibodies the difference in secretion may arise because the codon usage in some of the variable regions of heavy and light chains was optimized for *A. niger*. Koda and co-workers suggested that redesigning the primary DNA sequence of plant *Solanum tuberosum* alpha-glucan phosphorylase could be an effective method for improving heterologous protein production in *A. niger*.[24]

The aim of our work was to achieve secretion of human TNFα by the filamentous fungus *A. niger*. Several attempts have been made to obtain, improve or influence secretion. In this paper the major impact of codon usage and primary gene structure on secreting TNFα is discussed in detail.

2. Experimental

2.1. Construction of TNFα Expression Plasmids

Strains *Escherichia coli* JM109 or DH5α were used as hosts for molecular cloning. cDNA sequence TNFE encodes a synthetic gene, optimized for expression in *E. coli*, whereas TNFH is the cDNA for human TNFα.

pANGKT (GLA:$_{KEX}$:TNFE) is an expression plasmid for a glucoamylase-TNFE fusion protein having a site for *in vivo* cleavage with kexin protease.[25]

pANGT (GLA::TNFE) lacks a proteolytic cleavage site. It was constructed from pANGKT, which was cut with *Nhe*I and *Sna*BI restriction endonucleases. The smaller fragment was then replaced by a synthetic linker from two oligonucleotides, AI and AII. TNFE from the BBG4 plasmid (British Biotechnology) was inserted in the correct reading frame behind glucoamylase, the cleavage site was omitted.

pANXTNFH (GLA:$_{KEX}$:TNFH) is an expression plasmid for a glucoamylase-TNFH fusion protein harbouring a site for *in vivo* cleavage with kexin protease. TNFH was amplified from pE4 plasmid (ATCC) in a PCR reaction with NARXTNFH and HINDTNFH primers. It was inserted after *Nar*I and *Hind*III restriction in pAN56-2M plasmid.[26]

pANTNFH (GLA::TNFH) is as the former except without a proteolytic cleavage site. It was constructed as above, but with NARTNFN primer.

pANHETNFH (HisGLA:$_{EK}$:TNFH) is a plasmid for expressing fusion protein glucoamylase-TNFH, harbouring the His-tag and a site for *in vitro* cleavage by enterterokinase.[27]

pCODMUT1-5 plasmids (GLA:$_{KEX}$:TNFEm), with DNA mutations introduced having no influence on the amino acid sequence of TNFα, were prepared from pANGKT by Transformer™ Site-Directed Mutagenesis Kit (Clonetech). CODMUT1-5 oligonucleotides were used as mutagene and SEL-ECO as selection oligonucleotides.

pCODMUT6 (GLA:$_{KEX}$:TNFHm) is the same kind of plasmid as above, but CODMUT6 was used as mutagene and PANSEL2 as selection oligonucleotide on pANXTNFH.

pANXTNFH1 (GLA:$_{KEX}$:TNFH1) is as well an expression plasmid for a glucoamylase-TNFH fusion protein harbouring a site for *in vivo* cleavage with kexin protease, but the gene for human TNFα contains one intron (772 bp) (Figure 1). It was amplified from genomic DNA with NARXTNFH1 and HINDTNFH1 primers, and after restriction with *Nar*I and *HindIII* restriction endonucleases inserted in pAN56-2M plasmid.

pANXTNFH2 (GLA:$_{KEX}$:TNFH2) is the same as former except for the gene which contains two introns (959 bp) (Figure 1), amplification from genomic DNA was carried out with NARXTNFH2 and HINDTNFH1 primers.

pANHETNFH1 (HisGLA:$_{EK}$:TNFH1) is an expression plasmid for glucoamylase-TNFH fusion protein, harbouring the His-tag and a site for *in vitro* cleavage by enterterokinase, NARTNFEH1 forward primer was used for amplification the gene for human TNFα with one intron, PCR product was inserted in pAN56-MH plasmid.[27]

pANHETNFH2 (HisGLA:$_{EK}$:TNFH2) is as pAN HETNFH1, but the gene for human TNFα with two introns was amplified with NARTNFEH2 as a forward primer.

In all plasmids, selection marker *amdS* was introduced at the *Not*I restriction site. After amplification of the heterologous gene in a PCR reaction, the DNA sequence of the gene was verified by sequencing. Oligonucleotide primers sequences are presented in Supplemental table 1.

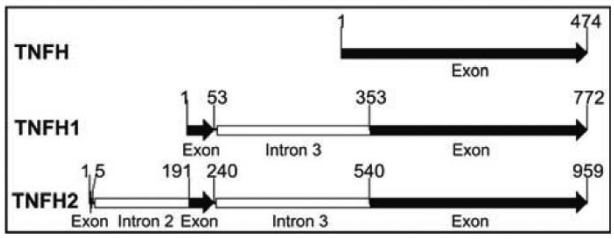

Figure 1. Schematic representation of different types of TNFα genes used for expression in *A. niger.*

2.2. Strain, Transformation Procedure and Growth Conditions

A protease deficient strain of *Aspergillus niger* AB1.13 (*cspA1, prt-13, pyrG1*)[28] (TNO, Zeist, NL) was used as a recipient for transformation. The protoplast formation and transformation procedure was carried out as described by Punt et al.[29] Transformants were selected on osmotically stabilized minimal medium plates with addition of 1 mM acetamide or acrylamide. For routine production, and for RNA isolation the mycelia were cultivated in Erlenmayer flasks with maltodextrin medium for 39 h.[26]

2.3. Northern Blot Analysis

Total RNA was isolated from powdered mycelia and Northern blot was carried out as described.[12] The DIG-TNF-probe was an equimolar mixture of both types of TNFα gene (size 0.5 kb), TNFE and TNFH. TNFE was amplified in a PCR reaction with DIG DNA Labelling Mix (Roche) from BBG4 plasmid with NARETNFE and HINDTNFE primers, while TNFH from pE4 plasmid with NARETNFH and HINDTNFH primers. Hybridization in DIG Easy Hib hybridization buffer (Roche) was performed together with DIG-GPD-probe as an internal control.

2.4. RT-PCR Reaction

RNA sample of GKT strain was treated with DNase I (Life Technologies) prior reverse transcription which was performed with the Reverse Transcription System (Promega) and AMV-reverse transcriptase. PCR reaction products were amplified with Titanium Taq DNA-polymerase (Clonetech) with RTTNF1 as forward and RTTNF201, RTTNF300, RTTNF350, RTTNF400 or RTTNF471 as reverse primers. Amplification with PEPC1 and PEPC2 primers was used as internal control.[30] PCR reaction products were analysed on DNALabChip 7500 by Agilent 2100 Bioanalyzer (Agilent Technologies).

2.5. Protein Analysis

SDS polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot were performed as described before.[25] To prepare intracellular samples, mycelium was filtered, frozen in liquid nitrogen and powdered. After disruption, the cells were resuspended in extraction buffer (50 mM Na_2HPO_4/NaH_2PO_4 pH, 7.0, 1 mM EDTA, 20 μM PMSF).

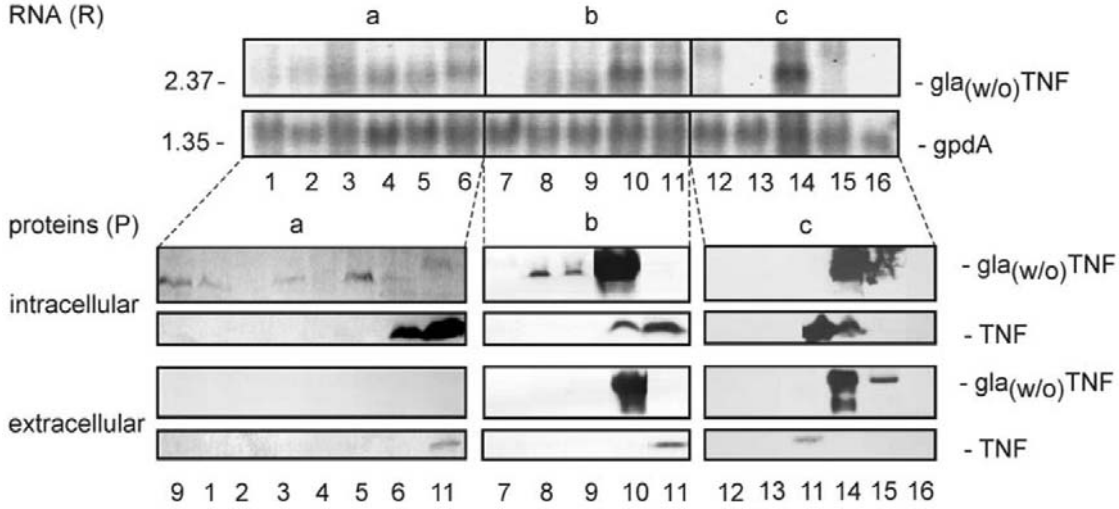

Figure 2. RNA (**R**) and proteins (**P**) from *A. niger* TNF transformants. Comparison of **a** – CODMUT strains; **b** – TNFE and TNFH strains and **c** – strains containing TNFH with introns. **1** – CODMUT1 (GLA:$_{KEX}$:TNFEm); **2** – CODMUT2 (GLA:$_{KEX}$:TNFEm); **3** – CODMUT3 (GLA:$_{KEX}$:TNFEm); **4** – CODMUT4 (GLA:$_{KEX}$:TNFEm); **5** – CODMUT5 (GLA:$_{KEX}$:TNFEm); **6** – CODMUT6 (GLA:$_{KEX}$:TNFHm); **7** – *A. niger* AB1.13 parent strain; **8** – GTA (GLA::TNFE); **9** – GKTA (GLA:$_{KEX}$:TNFE); **10** – TNFHA (GLA::TNFH); **11** – XTNFHA (GLA:$_{KEX}$:TNFH); **12** – XTNFH1 (GLA:$_{KEX}$:TNFH1); **13** – XTNFH2 (GLA:$_{KEX}$:TNFH2); **14** – HETNFH (HisGLA:$_{EK}$:TNFH); **15** – HETNFH1 (HisGLA:$_{EK}$:TNFH1); and **16** – HETNFH1 (HisGLA:$_{EK}$:TNFH2). Samples 9, 11 and 11 are used as reference values in protein blots (**P**) **a** and **c**, respectively. **gpdA** is a loading control, and **gla$_{(w/o)}$TNF** is glucoamylase-TNFα fusion mRNA with/without cleavage site. Molecular weights are indicated on the left side of the figure (kb).

2.6. RNA Secondary Structure Prediction

RNA secondary structure and corresponding stability were predicted for the TNFα coding region at 30 °C with the *mfold version 3.0* computer programme, and energy parameters of the Turner group version 2.3.[31,32] The algorithm calculated a structure with a minimal energy together with a population of other secondary structures with 5% higher energy of stabilization. The number of different associations of one base in this population is the *P-num* value; *ss-count* is the propensity of a base to be single stranded.

3. Results

3.1. Protein and RNA Analysis

In initial experiments, the synthetic cDNA optimized for expression in *E. coli* (TNFE) was used to express human TNFα in the filamentous fungus *Aspergillus niger*. Instead of secretion of TNFα into the medium as expected, a higher molecular weight protein was observed inside the cell (Figure 2 (P) samples 8 and 9). Samples were further analysed at the RNA level. Results of the Northern blot mRNA analysis (Figure 2 (R) samples 8 and 9) showed a rather weak and diffused signal.

The length of TNFE-mRNA of GKT strain was estimated by RT-PCR reaction, by amplification of fragments of gradually increasing size: 201, 300, 350 and 400 bp respectively (Figure 3 samples 1, 2, 3, 4). We succeeded in amplifying all the differently sized fragments, even the longest one of 471 bp (Figure 3 sample 5). The total size of the coding gene was 2384 bp, due to the GLA part. The authenticity of this fragment was verified after *Kpn*I restriction; two fragments of expected size (153 and 318 bp) were detected. The internal control is represented by a 200 bp band.

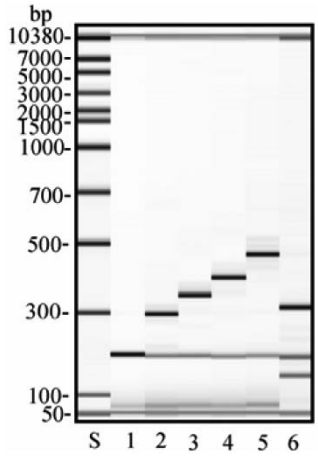

Figure 3. Determination of TNFE-mRNA length. RT-PCR reaction products were amplified from GKT cDNA, with forward RTTNF1 and different reverse primers: **1** – RTTNF201, **2** – RTTNF300, **3** – RTTNF350, **4** – RTTNF400, and **5** – RTTNF471. **6** – sample 5 after *Kpn*I restriction; **S** – DNA-leader; molecular standards are shown on the left side of the figure (bp).

3.2. Codon Usage

We compared the average codon usages of all genes for the fungus *A. niger*, *E. coli* and human, after complete sequences of their genomes were released (Figure 4). The codon usage of *A. niger* differs from that of *E. coli*, but human codon usage and that of *A. niger* are both GC rich. Typical codons from *A. niger* terminate in G or C (except Asp), furthermore, all codons that could terminate in G or C, end in C (except Leu). Average preferred codons show discrepancy for the following four amino acids: Arg, Asp, Ser and Val.

After replacement of TNFE by a natural, human cDNA (TNFH), TNFα cleaved from the fusion protein was detected in the culture medium (Figure 2 (P) samples 11). The two gene sequences are 76% identical at the level of base composition, but only 35% at the level of entire triplets. GC content at the third position of codons is 76% in TNFH and 53% in TNFE. Overall, TNFH contains 26% more GC base pairs. From 157 amino acids of TNFα, 90 in TNFH compared to 70 are encoded by average preferred codons in *A. niger*, and 28 compared to 33 by average least frequently used codons in the fungus.

After comparison of the primary sequence of TNFE with the natural human cDNA sequence, a cluster of seven less frequently used codons coding for residues from A^{109} to E^{116} in the last third of TNFE was noticed. Two contigous codons are among four not present in the glucoamylase gene (Figure 4), which is normally highly expressed under conditions suitable for heterologous gene expression. A whole set of mutants was generated by site-directed mutagenesis, where one codon (in CODMUT1 and CODMUT2 strains), two (in CODMUT3 strain) or five of the codons coding for residues from A^{109} to E^{116} in TNFE were exchanged from the less suitable bacterial to the more effective eukaryotic, as it is in TNFH or GLA (CODMUT4 and CODMUT5 strains). In contrast, one of the mutants (CODMUT6 strain) was prepared with the opposite exchange. We found that none of the mutations based on TNFE led to successful secretion of TNFα (Figure 2 (P) samples 1 to 5). In spite of the appearance of cleaved TNFα among the cell proteins of the CODMUT6 strain, no secreted TNFα was detected (Figure 2 (P) sample 6).

3.3. Prediction of mRNA Secondary Structure

Prediction of mRNA secondary structure and its resulting stability showed a difference in the energy of stabilization between TNFH (–185.0 kcal/mol) and TNFE-mRNA (–151.9 kcal/mol) (Table 1). In addition, the number of possible structures in a population with slightly higher energy of stabilization was lower in the case of TNFH (29) than for TNFE-mRNA (42). The prediction of the secondary structure of the TNFE was, particularly in the second half, less reliable than for TNFH-mRNA regar-

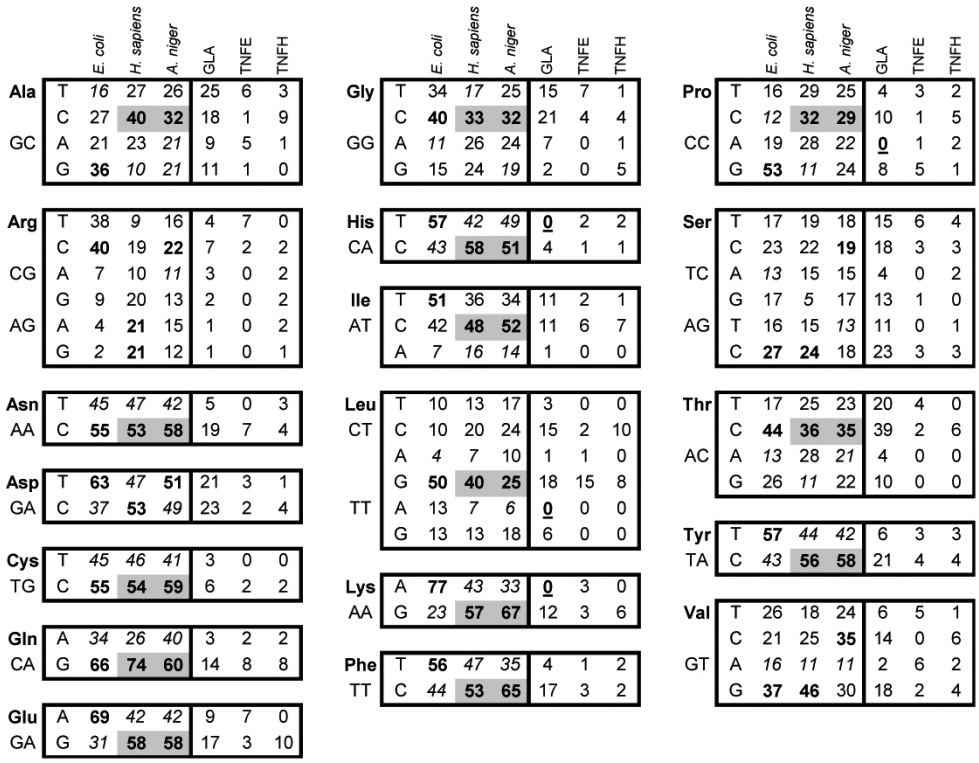

Figure 4. Average codon usage. Data (% per codon bias) are from CUTG (Codon Usage Tabulated from GenBank 156.0) from 4.331 CDS's *E. coli* and 89.533 CDS's *Homo sapiens*.[33] For 14.165 CDS's *A. niger*, data are from Pel et al.[34] Complete sequences of their genomes have been released. The least (*in italics*) and the most frequent (**in bold**) triplets for the same amino acid are indicated. Equal most frequent codons for *A. niger* and *H. sapiens* are marked in dark field.

ding the *P-num* value and the *ss*-number. Values for the mutants were between those for TNFE and TNFH, with increased reliability of prediction of their mRNA secondary structure in the region of residues A^{109} to E^{116}.

3.4 Inclusion of Introns

Our latest experiments elucidated the influence of introns of the human TNFα gene on secretion of TNFα protein in *A. niger*. The human TNFα gene contains three introns, one is present within the signal peptide encoding sequence (omitted in our studies) and the two within the TNFα coding region (Table 2). Predictions of RNA secondary structures suggested they might influence the primary transcript stability (Table 1), which could in turn influence heterologous expression of TNFα.

We prepared four different gene fusion constructs, containing one or both introns in the TNFα coding region.

Table 1. mRNA secondary structure prediction data.

Strain	ΔG	Structures	A^{109}	E^{110}	A^{111}	K^{112}	P^{113}	W^{114}	Y^{115}	E^{116}
GKT	–151.9	42	*gca*	*gaa*	*gcg*	*aaa*	*cca*	**tgg**	*tat*	*gaa*
CODMUT1	–152.0	38	*gca*	*gaa*	*gcg*	**aag**	*cca*	**tgg**	*tat*	*gaa*
CODMUT2	–154.0	42	*gca*	*gaa*	*gcg*	*aaa*	**ccc**	**tgg**	*tat*	*gaa*
CODMUT3	–154.1	40	*gca*	*gaa*	*gcg*	**aag**	**ccc**	**tgg**	*tat*	**gag**
CODMUT4	–154.8	37	gct	**gag**	**gcc**	**aag**	**ccc**	**tgg**	*tat*	**gag**
CODMUT5	–157.4	30	gct	**gag**	gct	**aag**	**ccc**	**tgg**	**tac**	**gag**
CODMUT6	–176.8	29	*gca*	*gaa*	*gcg*	*aaa*	*cca*	**tgg**	*tat*	*gaa*
XTNFH	–185.0	29	gct	**gag**	**gcc**	**aag**	**ccc**	**tgg**	*tat*	**gag**
TNFH1	–320.7	11								
TNFH2	–391.6	13								

The algorithm has calculated a structure with minimal free energy (ΔG) in kcal/mol for the TNFα coding region at 30 °C together with a population of other secondary structures (Structures) with 5% higher energy. Same sense mutations in a part of the TNFα gene are underlined. The least (*in italics*) and the most frequent codons (**in bold**) for the same amino acid residue as used in *A. niger* (Figure 4) are indicated.

Table 2. Intron structure of *A. niger* glucoamylase and human tumor necrosis factor (TNFα) genes

***A. niger* GLA**	**Intron A 75 bp**	ct/**gtatgt**ttcgagctcagatttagtatgagtgtgtcattgattgattga**tgctgac**tggcgtgtcgtttgttg**tag**
	Intron B 55 bp	tt/**gtatgt**tctccacccccttgcgtctgatctgtgacatatgt**agctgac**tggt**cag**
	Intron C 61 bp	tg/**gtgtgt**ttgttttatttttaaatttccaaagatgcgccagcag**agctaac**ccgcgatcg**cag**
	Intron D 58 bp	tg/**gtaagt**ctacgctagacaagcgctcatgttgacagagggtgcg**tactaac**agaag**tag**
	Intron E 169 bp	*ca/**gtacgt**catcaacctcctgtaccactcccaccgccgtggctgtgactttcgatctgacagctaccaccacctacggcgagaacatctac ctggtcggatcgatctctcagctgggtgactgggaaaccagcgacggcatag**c**tctgag**tgctgac**aagtacacttc**cag***
Human TNFα	***Intron 1 (Signal peptide) 597 bp***	*ag/**gtgagt**gcctggccagccttcatccactctcccacccaaggggaaatgagagacgcaagagagggagagagatgggatgggtgaa agatgtgcgctgatagggagggatgagagagaaaaaaacatggagaaagacggggatgcagaaagagatgtggcaagagatgggg aagagagagagagaaagatggagagacaggatgtctggcacatggaaggtgctcactaagtgtgtatggagtgaatgaatgaatgaa tgaatgaacaagcagatatataaataagatatggagacagatgtggggtgtgagaagagagatgggggaagaaacaagtgatatgaa taaagatggtgagacagaaagagcgggaaatatgacagctaaggagagagatgggggagataaggagagaagaagataggtgtct ggcacacagaagacactcagggaaagagctgttgaatgctggaaggtgaatacacagatgaatggagagagaaaaccagacacctca gggctaagagcgcaggccagacaggcagccagctgttcctcctttaagggtgactccctcga**tgttaac**ca**ttctccttctcccc**aa**cag***
	Intron 2 187 bp	ca/**gtaagt**gtctccaaacctctttcctaattctgggtttgggtttgggggtagggttagtaccggtatggaagcagtgggggaaatttaaagtttt ggtcttgggggaggatggatggaggtgaaagtagggggtattttctaggaagtttaagg**gtctcagctttttctttttctctctcctcttcag**
	Intron 3 301 bp	ag/**gtaaga**gctctgaggatgtgtcttggaacttggagggctaggatttggggattgaagcccggctgatggtaggcagaacttggagacaa tgtgagaaggactcgctgagctcaagggaagggtggaggaacagcacaggccttagtgggatactcagaacgtcatggccaggtgggatgt gggatgacagacagagaggacaggaaccggatgtggggtgggcagagctcgagggccaggatgtggagagtgaaccgacatggcca**ca ctgactctcctctcctctctcccctccctccag**

Consensus splice-site elements for a typical metazoan intron:[35] 5' Splice site (**guragu**); Branch point (**ynyurayy**); Polypyrimidine tract (y_{10-20}), 3' Splice site (**yag**); Pyrimidine (y); Purine (r); Any nucleotide (n); Two introns omitted from our studies are *in italics*.

Genes coding for glucoamylase-TNFH1 fusion proteins (of size 2685 bp) contain in total five introns: four from the GLA part and one from the TNFα part (intron 3). Genes coding for glucoamylase-TNFH2 fusion proteins (of size 2872 bp) contain in total six introns, the last two from the TNFα part (introns 2 and 3). Intron 3 from the human TNFα gene was correctly spliced in the fungus (Figure 2 (P) sample 16), as could be observed from successful secretion of the HisGLA:$_{EK}$:TNFH1 fusion protein. However, the same did not happen in the case of constructs containing both introns of the TNFα coding region.

4. Discussion

A gene fusion of TNFα with the shorter form of glucoamylase-G2 ($GlaA_{514}$) was constructed to express human TNFα in the filamentous fungus *Aspergillus niger*. To keep degradation of the protein as low as possible, a recipient host strain *A. niger* AB1.13 was transformed.[28] Although the most promising expression strategy was applied, instead of secretion of TNFα (TNFE coded) into the medium, a higher molecular weight protein was observed inside the cell, most probably the uncleaved fusion protein (Figure 2 (P) samples 8 and 9). Due to glycosylation of the GLA part of the fusion protein we could not determinate the exact protein size.

Results of the Northern blot RNA analysis suggested that the lack of secretion in the case of TNFE was already detectable at the mRNA level. Because a single band of the full length of TNFE-mRNA of GKT strain was estimated by RT-PCR, we concluded that a small amount of the full-sized and unstable mRNA was present, but was rapidly degraded, resulting in decreased translation and lack of secretion. Both results of mRNA secondary structure prediction were in keeping with the expected lower stability of the TNFE-mRNA. In case of xyloglucan endotransglucosylase from plant *Solanum tuberosum* expressed in *A. oryzae*, analysis specific mRNA revealed three bands of different sizes, only one of the predicted full length mRNA transcript, indicating incorrect processing of the specific cDNA at the pre-mRNA level. Authors suggested, that a high AT content and the occurrence of *A. oryzae* intron donor, acceptor, and branch point recognition sites resulted in erroneous intron interpretation of parts of the mRNA coding sequence.[36] This supposition may explain the difficulties generally experienced in expressing heterologous genes in filamentous fungi.

The lack of secretion was overcome by human TNFα cDNA (TNFH). The yield of cytotoxicaly active TNFα was 13 mgl^{-1} of the maltodextrin medium.[27] The N-terminus of TNFα resulting from cleavage of the fusion protein was not completely authentic. The shorter N-terminus could result either from the activity of other proteases, or from non-specific kexin processing, or both, as observed also by Punt et al.[37]

The codon usage signifies the average relative frequency of synonymous codons for one amino acid and is

strain specific. As each species has a different codon usage, expression of heterologous genes may be inhibited by inappropriate codons being present within the gene sequence. We compared the average codon usages of all genes for the fungus *A. niger*, *E. coli* and human, however, it is difficult to predict which difference is significant enough. Altering codon usage is not completely straightforward, with other factors having to be taken into account, such as mRNA stability and the possible presence of cryptic sequences within the heterologous gene. The extent of influence that codon usage exerts upon protein levels is still not completely known.[38]

Prediction of RNA secondary structure showed a difference in the energy of stabilisation between TNFE and TNFH-mRNA in favour of the latter. Also, the number of possible structures in a population of slightly higher energy of stabilisation was lower in the case of TNFH. Both numbers were in accordance with the expected lower stability of the mRNA in the case of TNFE gene expression, values for the TNFEm and TNFHm mutants were in between. The number of the different associations of one base in a population of other secondary structures with a slightly higher energy is the *P-num* value, while *ss*-number is the propensity of a base to be single stranded.[31,32] For RNA with a higher *P-num* value, the way to achieve a conformation associated with lower energy is expected to be more complicated and longer, and, in consequence, the action of RNases could be increased. Prediction was, particularly in the second half, less reliable for TNFE than for TNFH-mRNA regarding the *P-num* value and the *ss*-number (data not shown). For mouse TNFα (mTNF) cloned in bacterium *Streptomyces lividans*, no great difference in free energy of stabilization due to the different codon usage was detected, but the high producing strain had a low *P-num* value, and the low producing strain a high value.[39]

It appeared that in our experiment changing the codon usage and GC richness from the less suitable bacterial to the more effective eukaryotic was carried out in a too short region from A^{109} to $E^{116}$of TNFE to allow for successful secretion of TNFα into the medium. In the mutant with the opposite exchange, in spite of the appearance of cleaved TNFα among the cell proteins, no secreted TNFα was detected, probably due to lower production of this protein compared to TNFH. We estimated that by changing codons in a larger region, a more stable structure for the mRNA would result that could lead to more successful secretion of TNFα into the medium. When the influence of the codon usage on expression and secretion of the mTNF in *S. lividans* was studied, the authors exchanged five less frequently used codons at the N-terminus of the mTNF-gene to the bacterial ones. Unexpectedly, a sharp decrease in mTNF production was observed. They concluded that less frequently used codons are not responsible for lower heterologous protein expression, but the changes of the codon usage in a DNA sequence could cause changes of primary and secondary structures of mRNA, which would influence mRNA stability.[39]

We achieved successful expression of human cDNA (TNFH) in *A. niger*. The human encoded gene could thus be considered a good choice regarding codon usage, as an alternative to preparing a synthetic gene, which could be a relatively expensive approach. cDNA is usually used for expression in bacteria (e. g. *E. coli*), the question, however, remains as to the impact of introns on heterologous gene expression in eukcaryotes (e. g. fungi). Introns enhance almost every step of gene expression, and it was also found that human intron-containing genes have more stable mRNAs than intronless genes.[40]

Although the 5' and 3' splice sites and branch sites of fungal introns have a general consensus sequence, they differ subtly from metazoan consensus sequences in the respective regions. The number of nucleotides between the branch point A and the 3' splice site is within the range found for mammalian introns (11–40). Thus, the branch sites in fungal introns are positioned appropriately to function in splicing in a manner similar to that of branch sites in metazoan introns.[41] In our case, the distance between the branch point A and the 3' splice site in intron structure of *A. niger* GLA is even shorter for 2 out of 4 introns of our constructs. This observation is also confirmed by recent investigation of the introns of the model organism *Neurospora crassa*.[42]

On the other hand, the size of fungal introns of glucoamylase-TNFH1 fusion coding genes was 55–75 bp, which fits well the narrow peak of size distribution in the three fungi *A. nidulans*, *N. crassa*, and *C. neoformans*. Human coded introns from our constructs are longer, 187 and 301 bp respectively. This length difference is accounted for by the distance from the 5' splice site to the branch point. The polypyrimidine tract is absent between the 3' splice site and the branch point of *A. niger* GLA introns, but is present within both human TNFα introns, which is the most striking difference observed between the fungal and mammalian introns absent in half of the intron population.[41]

Evidently, there are significant similarities between introns of vertebrates and fungi. There are, however, some important differences that have an impact on the mechanisms used for intron splicing, and which influence heterologous secretion in *A. niger* as well. One intron of the TNFα coding region was correctly spliced in *A. niger*, as could be observed from successful secretion of the TNFH1 fusion protein. However, due to acumulation of intrinsc intron characteristics, the same did not happen (fast enough) in the case of constructs containing two introns. For further experiments, we propose introduction of avarege fungal intron(s) that might be beneficial for enhanced protein production and secretion. For example, the green fluorescent protein gene functioned as a reporter of gene expression in *Phanerochaete chrysosporium* only when an intron in the form of the first intron from *gpd* ge-

ne (55 bp) was inserted.[43] Introns were necessary for mRNA accumulation in *Schizophyllum commune*. When a 50 bp artificial intron containing the consensus splice and branch sites of *S. commune* introns, in addition to randomly-generated sequences, was introduced in the correct orientation into the intronless SC3 transcriptional unit, the accumulation of SC3 mRNA was restored.[44]

5. Conclusions

In conclusion, we achieved successful secretion of human TNFα by the filamentous fungus *Aspergillus niger*. The most important factor influencing secretion was variation of the primary sequence of the heterologous gene. Non-optimized codon usage in the TNFα gene constructed for use in bacteria expression influenced badly expression of TNFα in *A. niger* at the mRNA level, provoked by unstable mRNA with a low transcription rate. The observation of the drastic effect of codon usage on secretion of TNFα was supported by secondary structure prediction data. We believe that changes of the codon usage or GC richness which altered chemical features were made in too short a region of the TNFα gene to allow successful secretion of TNFα into the medium. In fact, by changing codons over a larger region, i.e. by replacing the synthetic gene optimized for expression in *E. coli* (TNFE) by a natural, human cDNA (TNFH), more stable mRNA structures were produced, which led to more successful secretion of TNFα into the medium. On the basis of the high impact of codon usage on successful heterologous secretion observed in our experiments, we suggest that optimization of a heterologous gene should always be the first step in constructing an efficient fungal secretion system. Predictions of RNA secondary structures suggested that introns within the gene coding region may have influence on primary transcript stability, which could influence heterologous expression. In this context, strategy of gene fusion constructs, containing heterologous/homologous intron(s) seem promissing.

6. Acknowledgements

The work was sponsored by the research grant P1-0518-0104 provided by the Ministry of Education, Science and Sport of Slovenia. The authors thank Jelka Lenarčič, Nataša Lileg Tašler and Ana Jesenko for their excelent technical assistance and Dr. Roger Pain for his review of this paper. To Ljerka Lah!

7. References

1. F. J. Lejeune, C. Ruegg, D. Lienard, *Curr. Opin. Immunol.* **1998**, *10*, 573–580.
2. L. G. Frenken, J. G. Hessing, C. A. M. J. J. van den Hondel, C. T. Verrips, *Res. Immunol.* **1998**, *149*, 589–599.
3. R. J. Gouka, P. J. Punt , C.A.M.J.J. van den Hondel, *Appl. Microbiol. Biotechnol.* **1997,** *47*, 1–11.
4. D. B. Archer, J. F. Peberdy, *Crit. Rev. Biotechnol.* **1997**, *17*, 273–306.
5. W. E. Hintz, I. Kalsner, E. Plawinski, Z. M. Guo, P. A. Lagosky, *Can. J. Bot.* **1995**, *73*, S876–S884.
6. J. S. Kruszewska, *Acta biochim. Pol.* **1999**, *46*, 181–195.
7. D. A. Mackenzie, D. J. Jeenes, N. J. Belshaw, D. B. Archer *J. Gen. Microbiol.* **1993**, *139*, 2295–2307.
8. J. P. T. W. Van Den Hombergh, P. J. I. Van de Vondervoort, L. Fraissinet-Tachet, J. Visser, *Trends. Biotech.* **1997**, *15*, 256–263.
9. J. C. Verdoes, P. J. Punt, C. A. M. J. J. van den Hondel, *Appl. Microbiol. Biotechnol.* **1995**, *43*, 195–205.
10. K. M. Nevalainen, V. S. Te'o, P. L. Bergquist. *Trends Biotechnol.* **2005**, *23*, 468–474.
11. S. Mestrić, P. J. Punt, R. Valinger, C. A. M. J. J. Van den Hondel, *Fungal. Genet. Newsl.* **1996**, *43B*, 25.
12. N. Kraševec, C. A. M. J. J. van den Hondel, R. Komel, *Pflugers arch. Eur. J. Physiol.* **2000**, *440*, R83–R85.
13. M. G. Wiebe, A. Karandikar, G. D. Robson, A. P. Trinci, J. L. Candia, S. Trappe, G. Wallis, U. Rinas, P. M. Derkx, S. M. Madrid, H. Sisniega, I. Faus, R. Montijn, C. A. M. J. J. van den Hondel, P. J. Punt, *Biotechnol. Bioeng.* **2001,** *76*, 164–174.
14. P. J. Punt, N. van Biezen, A. Conesa, A. Albers, J. Mangnus, C. A. M. J. J. van den Hondel, *Trends Biotechnol.* **2002,** *20*, 200–206.
15. M. Ward, C. Lin, D. C. Victoria, B. P. Fox, J. A. Fox, D. L. Wong, H. J. Meerman, J. P. Pucci, R. B. Fong, M. H. Heng, N. Tsurushita, C. Gieswein, M. Park, H. Wang, *Appl. Environ Microbiol.* **2004**, *70*, 2567–2576.
16. R. J. Gouka, P. J. Punt, J. G. M. Hessing, C. A. M. J. J. van den Hondel, *Appl. Environ. Microbiol.* **1996**, *62*, 1951–1957.
17. I. A. Van Gemeren, A. Beijersbergen, C. A. M. J. J. van den Hondel, C. T. Verrips, *Environ. Microbiol.* **1998**, *64*, 2794–2799.
18. I. Faus, C. Patino, J. L. del Rio, C. del Moral, H. S. Barroso, V. Rubio, *Biochem. Biophys. Res. Commun.* **1996**, *229*, 121–127.
19. I. Faus, C. del Moral, N. Adroer, J. L. del Rio, C. Patino, H. Sisniega, C.Casas, J. Blade, V.Rubio, *Appl. Microbiol. Biotechnol.* **1998**, *49*, 393–398.
20. B. Cormack, *Curr. Opin. Microbiol.* **1998**, *1*, 406–410.
21. D. Siedenberg, S. Mestrić, M. Ganzlin, M. Schmidt, P. J. Punt, C. A. M. J. J. van den Hondel, U. Rinas, *Biotechnol Prog.* **1999**, *15*, 43–50.
22. V. S. Te'o, A. E. Cziferszky, P. L. Bergquist, K. M. Nevalainen, *FEMS Microbiol. Lett.* **2000**, *190*, 13–19.
23. R. E. Cardoza, S. Gutierrez, N. Ortega, A. Colina, J. Casqueiro, J. F. Martin, *Biotechnol. Bioeng.* **2003,** *83,* 249–259.
24. A. Koda, T. Bogaki, T. Minetoki, M. Hirotsune, *J. Biosci. Bioeng.* **2005**, *100*, 531–537.
25. N. Kraševec, C. A. M. J. J. van den Hondel, R. Komel, *Pflu-*

gers arch. Eur. J. Physiol. **2000**, *439*, R84–R86.
26. M. Svetina, N. Kraševec, V. Gaberc-Porekar, R. Komel, *J. Biotechnol.* **2000**, *76*, 245–251.
27. N. Kraševec, M. Svetina, V. Gaberc-Porekar, V. Menart, R. Komel, *Food Technol. Biotechnol.* **2003**, *41*, 345–351.
28. I. E. Mattern, J. M. van Noort, P. van den Berg, D. B. Archer, I. N. Roberts, C. A. M. J. J. van den Hondel, *Mol. Gen. Genet.* **1992**, *234*, 332–336.
29. P. J. Punt, C. A. M. J. J. van den Hondel, *Meth. Enzymol.* **1992**, *216*, 447–457.
30. M. Benčina, M. Legiša, *Biotechnol. Techniques* **1999**, *13*, 865–869.
31. M. Zuker, *Nucleic Acids Res.* **2003**, *31*, 3406–3415.
32. A. E. Walter, D. H. Turner, J. Kim, M. H. Lyttle, P. Muller, D. H. Mathews, M. Zuker, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9218–9222.
33. Y. Nakamura, T. Gojobori, T. Ikemura, *Nucleic acids Res.* **2000**, *28*, 292.
34. H. J. Pel, J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann, P. J. Schaap, G. Turner, R. P. de Vries, R. Albang, K. Albermann, M. R. Andersen, J. D. Bendtsen, J. A. E. Benen, M. van den Berg, S. Breestraat, M. X. Caddick, R. Contreras, M. Cornell, P. M. Coutinho, E. G. J. Danchin, A. J. M. Debets, P. Dekker, P. W. M. van Dijck, A. van Dijk, L. Dijkhuizen, A. J. M. Driessen, C. d'Enfert, S. Geysens, C. Goosen, G. S. P. Groot, P. W. J. de Groot, T. Guillemette, B. Henrissat, M. Herweijer, J. P. T. W. van den Hombergh, C. A. M. J. J. van den Hondel, R. T. J. M. van der Heijden, R. M. van der Kaaij, F. M. Klis, H. J. Kools, C. P. Kubicek, P. A. van Kuyk, J. Lauber, X. Lu, M. J. E. C. van der Maarel, R. Meulenberg, H. Menke, M. A. Mortimer, J. Nielsen, S. G. Oliver, M. Olsthoorn, K. Pal, N. N. M. E. van Peij, A. F. J. Ram, U. Rinas, J. A. Roubos, C. M. J. Sagt, M. Schmoll, J. Sun, D. Ussery, J. Varga, W. Vervecken, P. J. J. van de Vondervoort, H. Wedler, H. A. B. Wösten, A.-P. Zeng, A. J. J. van Ooyen, J. Visser, H. Stam *Nat. Biotechnol.* **2007**, *25*, 221–231.
35. C. W. Smith, J. Valcarcel, *Trends Biochem. Sci.* **2000,** 25, 381–388.
36. T. Hamann, L. Lange. J. Biotechnol. **2006**, *126* 265–276.
37. P. J. Punt, A. Drint-Kuijvenhoven, B. C. Lokman, J. A. Spencer, D. Jeenes, D. A. Archer, C. A. M. J. J. van den Hondel, *J. Biotechnol.* **2003**, *106,* 23–32.
38. G. Nelson, O. Kozlova-Zwinderman, A. J. Collis, M. R. Knight, J. R. Fincham, C. P. Stanger, A. Renwick, J. G. Hessing, P. J. Punt, C. A. M. J. J. van den Hondel, N. D. Read. *Mol Microbiol.* **2004**, *52*, 1437–1450.
39. E. Lammertyn, L. Van Mellaert, A. P. Bijnens, B. Joris, J. Anne, *Mol. Gen. Genet.* **1996**, *250*, 223–229.
40. H.-F. Wang, L. Feng, D.-K. Niu. *Biochem. biophys. res. commun.* **2007**, *354*, 203–208.
41. D. M. Kupfer, S. D. Drabenstot, K.L. Buchanan, H. Lai, H. Zhu, D. W. Dyer, B. A. Roe, J. W. Murphy. Eukaryot. Cell **2004**, *5*, 1088–100.
42. K. L. Henscheid, R. B. Voelker, J. A. Berglund. *Biochemistry* **2008,** *47,* 449–459.
43. B. Ma, M. B. Mayfield, M. H. Gold. *Appl. Environ. Microbiol.* **2001,** *67*, 948–955.
44. L. G. Lugones, K. Scholtmeijer, R. Klootwijk, J. G. Wessels. *Mol. Microbiol.* **1999**, *32*, 681–689.

Povzetek

Proučevali smo izločanje dejavnika tumorske nekroze alfa (TNFα) iz nitaste glive *Aspergillus niger*. V primeru fuzije glukoamilaznega gena s TNFα, prilagojenim za izražanje v bakteriji *E. coli* (TNFE), smo zasledili TNFα samo v celicah, v obliki neodcepljene fuzijske beljakovine. Neoptimalna uporaba kodonov je neugodno vplivala na nastajanje mRNA TNFα. Področje z več neugodnimi kodoni smo s pomočjo mestno specifične mutageneze prilagodili za evkariontsko uporabo z vnosom kodonov, ki so sicer navzoči v humanem TNFα (TNFH) oziroma v glivni glukoamilazi, vendar izboljšanja nismo zasledili. TNFα se je izločil v gojišče šele, ko smo TNFE v celoti zamenjali s humanim TNFH. Če smo kodone v obravnavanem področju gena TNFH spremenili nazaj v neugodne, je bilo izločanje izničeno. Na osnovi velikega vpliva rabe kodonov, ki smo ga opazili v naših poskusih, ugotavljamo, da je humana cDNA primerna za izražanje tujega gena v *A. niger*, saj je kod tujega gena dovolj podoben kodu gostitelja. Metode napovedovnja sekundarne strukture mRNA so se izkazale kot uporabne za pripravo genskih konstruktov, vključno tistih s heterolognimi introni.

Supplemental table 1: Oligonucleotide primers

Name		Sequence
AI		CTAGCAAGACCAGCACCAGTACGTCATCAACCTCCTGTACCACTCCCACGGCGGTAC
AII		GTACCGCCATGGGAGTGGTACAGGAGGTTGATGACGTACTGGTGCTGGTCTTG
CODMUT1	f	GGGTGCAGAAGCGAAGCCATGGTATGAACCG
CODMUT2	f	GAAGCGAAACCCTGGTATGAACCG
CODMUT3	f	GCAGAAGCGAAGCCCTGGTATGAACCG
CODMUT4	f	CCCGAGGGTGCTGAGGCCAAGCCCTGGTATGAGCCGATCTACCTGG
CODMUT5	f	CCCGAGGGTGCTGAGGCTAAGCCCTGGTACGAGCCGATCTACCTGG
CODMUT6	f	CCCAGAGGGGGCAGAAGCGAAACCATGGTATGAACCCATCTATCTGG
HINDTNFE	r	GACCT***AAGCTT***TTATTATTACAGTGCGATAA
HINDTNFH	r	CTC***AAGCTT***TTATCACAGGGCAATGATCC
HINDTNFH1	r	CTC***AAGCTT***TTATCACAGGGCAATGATCCCAAAGTAGACC
NARETNFE	f	CGAC***GGCGCC***GGATGACGATGACAAGGTACGTAGCTCCTCTCGC
NARETNFH	f	CGAC***GGCGCC***GGATGACGATGACAAGGTCAGATCATCTTCTCGA
NARETNFH1	f	CGAC***GGCGCC***GGATGACGATGACAAGGTCAGATCATCTTCTCGAACCCCGAGTG
NARETNFH2	f	CGAC***GGCGCC***GGATGACGATGACAAGGTCAGTAAGTGTCTCCAAACCTCTTTCC
NARTNFH	f	CTC***GGCGCC***TGTCAGATCATCTTCTCGA
NARXTNFH	f	CTC***GGCGCC***TAATGTGATTTCCAAGCGTGTCAGATCATCTTCTCGA
NARXTNFH1	f	CTC***GGCGCC***TAATGTGATTTCCAAGCGTGTCAGATCATCTTCTCGAACCCCGAGTG
NARXTNFH2	f	CTC***GGCGCC***TAATGTGATTTCCAAGCGTGTCAGTAAGTGTCTCCAAACCTCTTTCC
PANSEL2		AATGGTTTCTTAGACGTCAGGTGGCAC
RTTNF1	f	GTACGTAGCTCCTCTCGCACTC
RTTNF201	r	CTGACCCTTGAACAGTACTTGAG
RTTNF300	r	CGGAGACTTGATAGCGCTCAG
RTTNF350	r	GGTTCATACCATGGTTTCGC
RTTNF400	r	CGGACAGACGGTCACCTTTC
RTTNF471	r	CAGTGCGATAATACCGAAGTACAC
SEL-ECO		GGTTTCTTAGATATCAGGTGGCAC

Forward (f), reverse (r), *Nar*I or *Hind*III restriction site is underlined