



Entelechon Catalog

Content (interactive)

perfectSequencing

perfectGene

perfectBiology

perfectProtein

perfectSilico

Technical Bulletin and Appendix

Table of Contents (interactive)

perfectSequencing

Versatile DNA analysis at your command
Sequencing Service

perfectGene - Gene synthesis and more

Gene Synthesis
Leto gene optimization
DNA libraries
Subcloning
Vector construction
Mutation Service
PCR

perfectBiology – Tailored solutions for your research

Oligonucleotide synthesis
Site-directed mutations
DNA libraries
PCR
Plasmid Upscaling Service
Subcloning

perfectProtein – Integrated solutions for protein biochemistry

Protein expression
Screening of expression clones
Protein purification
Protein characterization
Quantitative proteomics
Synthetic peptides
Antibodies

perfectSilico

Datamining
Database setup
Software Development
Microarray oligonucleotide design
Custom bioinformatics solutions
Clustering and homology analysis

Technical Bulletin and Appendix



perfectSequencing

Solutions for DNA quality
control and analysis





perfectSequencing

Dideoxy technique for DNA sequencing

Frederick Sanger

1980, Stockholm, Sweden: It is Sanger's second trip to Stockholm, it is his second Nobel Prize and for the second time he is honoured for sequencing. In 1958 the scientist analysed the amino acid sequence of insulin. 22 years later Sanger is the fourth Nobel laureate after Marie Curie, Linus Pauling and John Bardeen, who receives the coveted prize a second time (together with Walter Gilbert and Paul Berg). He just changed the object of his research from proteins to DNA and created the chain termination method (also called the Sanger method) of DNA sequencing.

In 1978 he succeeded in sequencing the genome of the phage phi-X174 – the first fully sequenced DNA-based genome, consisting of 5386 nucleotides.

When comparing this size with today's high throughput methods and the sequencing of the human genome, it is hard to imagine how formidable a task the sequencing of this tiny viral genome in the seventies really was.

Even today Sanger's method is the basis for a lot of sequencing that is performed worldwide. His "dideoxies" still work! Sanger is strictly opposed to commercialising genetic information derived from sequencing projects, he advocates for open access to all these data.

Today it becomes more and more obvious just how fantastic the world is to which Sanger's breakthrough opened the door. And even in a world of the 1000 dollars genome, Sanger's feat will not be forgotten.



perfectSequencing

Sequencing the World is not Enough

*The name's Craig Venter, and it has become a trademark among biologists. The American scientist and entrepreneur with a preference for challenging projects has been nicknamed "lord of the genes". His treasure: A shotgun technique – and the money – to sequence them all: His private, non-profit "Institute for Genomic Research" (TIGR) sequenced the whole genome of *Haemophilus influenzae* in 1995:*

*Soon the data for *Mycoplasma* and *Methanococcus* followed this first complete DNA sequence of a freelifing organism. Venter aimed higher and participated in the Human Genome Project (HGP).*

By now, he has also analyzed genes of mice and monkeys, of the fruit fly, the whole Sargasso Sea, the air of midtown New York City, samples of Australia's soil, and last but not least his own genome.

The shotgun sequencing approach is equally efficient and effective: Genomic DNA is cut into small fragments and cloned into vectors. Then every single DNA piece is sequenced from both ends – and finally high performance computing is thrown at the sequence fragments to assemble contiguous chromosome sequences.

With this technique, Venter won the sequencing race with the publicly funded HGP to read the human genetic code: The HGP scientists decided to walk over the human chromosomes step by step instead of parallelizing.

In 2000 Venter announced the mapping of the human genome – and started to turn genome research into a commercial endeavour: He patented about 6.500 genes of interest for the pharmaceutical industry.

Venter published his own personal genome in the PLoS journal: The lord of the genes is the first person in history to gaze at his own genome consisting of about six billion letters – maybe the starting shot for personal genomics for everybody?

However, the genome is not the end of the story. As Venter observed himself: "We can get very few yes/no answers out of our genomes. While genetics influence every aspect of our existence, including behavior and personalities, there's no way to truly tell right now what's caused by our genes and what comes from the environment".



perfectSequencing

Versatile DNA analysis at your command

Entelechon provides a wide range of DNA analysis services. We take pride in our customer-oriented workflow which allows for the adaptation of DNA sequencing protocols, preparation and documentation to the client's requirements.

Entelechon's services include standard capillary sequencing, but also primer walking, mutation sequencing and primer design.

Our DNA sequencing and bioinformatics units work closely together for complex projects. This and an extremely scalable workflow enable us to cover sequencing projects of any size and scope.

Entelechon has successfully conducted large-scale clone library projects in the past. A range of specifically adapted protocols cover non-standard sequencing problems, such as templates with an unusually high GC content.

Results are conveniently provided in electronic format, and our molecular biology experts are ready to provide in-depth tech support. DNA sequencing is dependent on the prime condition of samples. If something goes wrong, we go to great lengths to find out and explain what caused the problem - be it a wrong concentration of the template, an incompatible buffer composition or something else entirely.

Documentation

- In FASTA and text format as well as ABI chromatograms
- Manually edited sequencing chromatograms
- Data available to download from a secure and password protected on-line account or sent to your e-mail address
- Data provided within 1-2 work days after receipt of the samples



perfectSequencing

Sequencing Service

SimpleRun

Sequencing of plasmids and PCR products up to 600 nt in top quality

SimpleRun

SEQ-001

DoubleRun

Sequencing of sense and antisense DNA of plasmids and PCR products up to 600 nt in top quality

DoubleRun

SEQ-002

FlankingRun

Sequencing of flanking regions of vector inserts with up to 600 nt in top quality

FlankingRun

SEQ-003

PrimerWalk

Single strand sequencing of plasmids, starting at an initiator site, including primer design and synthesis for the newly determined sequence and sequential delivery of the results.

PrimerWalk

SEQ-004

MutationSequencing

Sequencing of plasmids and PCR products at the mutation site in top quality

MutationSequencing

SEQ-005



perfectSequencing

Free services

- Storage of DNA and primers for future sequencing projects (8 weeks and more after consultation)
- Standard primers
- Hotline support (9 a.m. - 6 p.m., CET)

Sequencing account and courier service

For clients with a minimum of 100 sequencing runs we offer the possibility to establish a sequencing account with attractive conditions. Sequencing runs are then deducted from a previously established credit note, making the transaction easy and seamless.

For sequencing customers located within Regensburg, we maintain a pick-up-service free of charge in the afternoon. Your sequences will be processed on the same evening.

Free primers

For your sequencing orders we have primers for standard vectors on stock free of charge, for instance M13for/rev, T3/T7, Sp6, T7term and BGHrev. For more free standard primers, please consult our website at www.entelechon.com/sequencing.

Primer and oligo design

We provide in silico design of oligonucleotides for microarrays, hybridization experiments, PCR as well as for sequencing. If you require large-scale oligonucleotide design, please contact us for details.



perfectSequencing

How to order

If you are interested in Entelechon's DNA sequencing services, the next step is easy:

Call our sequencing specialist

+49 (9405) 96 999 10

or send an inquiry to

perfectSequencing@entelechon.com

Or visit www.entelechon.com where you will find additional information.



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com





perfectGene

Gene synthesis
Gene optimization
Gene libraries and randomizations
Custom subcloning
Vector construction
Site-directed mutagenesis
PCR



perfectGene

Har Gobind Khorana: The first artificial gene

*27th August 1976:
Khorana and a team of 24 re-
searchers finally manage to get the
first entirely artificial gene to work:
A tyrosine-specific, bacterial tRNA.
Although they finish synthesizing
this piece of DNA in 1973, it should
take another 3 years until the tRNA
gene fulfills its biological task in an
E.coli mutant that originally lacks
the genetic information for this
very gene.*

*Together with Robert W. Hol-
ley and Marshall W. Nirenberg,
Har G. Khorana was awarded the
Nobel Prize in Medicine in 1968 -
not for arranging A, G, C and T into
the spiral ladder of a DNA molecule
via advanced and labourous chem-
istry, but for establishing that the
mother of all codes is spelled out in
three-letter syllables.*

*It takes the Indian-American Nobel
prize laureate more than three
years to build a gene consisting of
207 base pairs in the 1970s. Today
Khorana's monumental feat can be
replicated in hours.*

*In the same year in which Khorana
announces his successful synthesis
the world's first commercial bio-
tech company "Genentech" invents
an automated method of synthe-
sizing whole genes: Artificial genes
for everyone – you just need to fax
the desired genetic sequence!*

*It took another 25 years before
commercial gene synthesis really
takes off, but it has since helped to
revolutionize the way we ma-
nipulate genomes, opening up the
exciting field of synthetic biology!*



perfectGene

Modern Animal Farm: Pigs glowing in the dark?

Everybody knows that pigs come in trademark pink, but what's their color at night - gray? Not any longer! An ominously fluorescent green is the new in-color for animals of any sort!

*In the 1960s the Japanese scientist Shimomura started studying the bioluminescence ability of the jellyfish *Aequorea*— and discovered the today well-known “Green Fluorescent Protein” (GFP). Other investigators soon noticed that this little protein is well-suited for labelling proteins of interest: Just attach the GFP sequence to a gene of your choice and all cells that are expressing this modified protein will shine green when treated with blue light. “Taiwan bred pigs that glow in the dark!” was a striking headline in 2006.*

But others have already created pigs with a greenish-yellow snout and trotters some years ago – just by treating swine embryos with

the genetic information for GFP. This experiment demonstrated that it is possible to create a transgenic clone of a mammal.

*In 2006, the same team produced another transgenic pig: This time they added the fat-1 gene to the embryo. The gene product of fat-1 is an enzyme that converts the “bad” omega-6 fatty acids into “healthy” omega-3 fatty acids. Professor Prather's team transfected special fibroblasts with the “humanized” version of the *C. elegans*-derived fat-1 gene. The result: Pink pigs with enhanced levels of omega-3 fatty acids in their muscles.*

For now, the obvious application - healthy pork chops - is not available in the market, but the experiment shows just how far genetic engineering has come along.



perfectGene

Gene synthesis

Entelechon has been in the business of custom gene synthesis for over ten years. We have developed a highly automated and optimized workflow for the assembly of genes of arbitrary length and composition.

Entelechon has created a proprietary software application for the optimization of gene expression, called Leto. Based on a genetic algorithm, this program is capable of optimizing a multitude of parameters which have an impact on gene expression. This includes, for instance, mRNA secondary structure, codon usage, and mRNA destabilizing motifs.

Entelechon's tech support team consists of expert molecular biologists who can advise our clients on the best design of a gene sequence for a given purpose.

In addition to individual genes, we have vast expertise in creating randomized gene libraries, based on the incorporation of wobble nucleotides at strategic positions. By combining bioinformatics solutions for the design of libraries with advanced protocols for the synthesis of randomized genes, we can ensure a high quality of the resulting libraries. We put strong emphasis on ensuring an optimal match with the desired nucleotide distribution, and we take great care to preserve full coverage of the theoretical variations.



perfectGene

perfectGene - Standard service

- Available in any length for standard genes (GC-content of 40-60%, no repeats, non-toxic for humans or E. coli, not on the export control list)
- Several different standard vectors available for free
- Free PCR primers available
- LETO gene optimization available on request

Standard gene synthesis

GS-001

perfectGene - Premium service

- Shortest production times and fast delivery for urgent orders
- Available in any length for standard genes (GC-content of 40-60%, no repeats, non-toxic for humans or E. coli, not on the export control list)
- Several different standard vectors available for free
- Free PCR primers available
- LETO gene optimization included

Premium gene synthesis

GS-002

perfectGene - Complex genes

- High AT/GC content, long repeats, and secondary structures possible
- Proprietary protocols tailored to the gene sequence - based on 10 years of expertise in gene synthesis
- Including personal consultancy to find the optimal expression system, to check your sequence in advance, and to design your perfectGene in silico
- Available for genes in any length
- LETO gene optimization available

Complex gene synthesis

GS-003



Leto gene optimization

The Leto gene optimization software package has been developed at the Entelechon bioinformatics unit over more than eight years. Leto uses a genetic algorithm which allows for a multi-objective optimization of gene sequences for maximum expression yields.

Expertise based on many years of gene optimization and customer feedback have continuously been incorporated into this software. Leto's advanced optimization algorithm allows it to search the vast space of possible codon permutations for a DNA sequence that optimally meets a number of important criteria, such as codon usage, mRNA secondary structure, avoidance of splice sites or mRNA stability.

Leto benefits include:

- Set of 15 optimization parameters with a direct impact on the expression yield
- Up-to-date and species-specific splice site prediction
- Simultaneous optimization for multiple target organisms
- Implementation of custom parameters or requirements
- Detailed documentation and consulting available

Leto optimization

GS-004



perfectGene

DNA libraries

For the construction of gene libraries, we can introduce randomized nucleotide positions during gene synthesis.

You are completely free to choose any mixture of nucleotides at any given position, including a uniform distribution of all four nucleotides or asymmetric mixes of two, three or four nucleotides.

Please note that DNA synthesis is not entirely error-free. The end product will contain small amounts of mismatches at the non-randomized positions. The exact ratio of erroneous versus correct products depends on a number of factors such as gene length and sequence composition and cannot be predicted a priori. However, Entelechon's strict quality control and tightly controlled synthesis process ensure the highest possible quality.

All libraries can be delivered as non-ligated PCR products or cloned into a standard or custom vector.

For further information on randomized DNA libraries please see our technical bulletin I.

DNA library

BIO-002



Subcloning

We can clone fragments provided by the client or produced at Entelechon into a range of standard target vectors as well as into custom vectors.

Examples of available vectors are:

Cloning vectors	Expression vectors
pSC-A	pET-vectors
pPCR-Script-Amp	pGEX-4T-1, pGEX-6p-1
pCR4-TOPO	pEGFP-C1
pUC18, pUC19	pIRES-vectors
pBluescriptII	pCDNA3.1
	pMAL-c2

Please inquire for other vector options and for subcloning into a vector provided by you.

Subcloning in-house vector BIO-007-001

Subcloning customer vector BIO-007-002



perfectGene

Vector construction

Building on the gene synthesis service, Entelechon performs vector construction projects according to a wide range of specifications. Choose from a range of target organisms, resistance and selection markers, and template designs.

Entelechon will provide in-depth consulting on the final specification and feasibility of a specific vector design. If you are unsure about requirements, we can add a design phase completed with test designs and performance evaluation.

Entelechon's DNA analysis and QC lab will provide full and detailed documentation of the final vector construct according to your requirements.

- Includes free project consultation, planning, and monitoring
- De novo synthesis of vector elements (promoters, resistance markers, etc.)
- Generation of vector components via PCR
- Stringent quality control
- Upscaling of delivered DNA quantity to the mg range, production under GMP conditions available

Vector construction

GS-005



perfectGene

Mutation service

Our mutation service allows to introduce one or more site-specific mutations into an existing DNA fragment, including insertions and deletions.

This is useful for a wide range of applications, such as functionality studies, interaction studies, knock-outs, and splice-site studies. Sometimes adapting an existing gene by site-directed mutagenesis is more efficient and faster than complete de novo synthesis.

Site-directed mutagenesis

BIO-001

PCR

Although PCR is a staple technique in most molecular biology labs, it still has its pitfalls. Depending on the desired application, PCR setup and development of a robust PCR protocol can be tricky.

Entelechon's advanced molecular biology lab can develop virtually any PCR application. Our long track record of successful implementations of PCR-based protocols helps us to attack very challenging problems, such as the development of multiplex PCR assays that simultaneously and semiquantitatively detect 24 targets.

Although by today's standards a very traditional method, PCR can still be the most time-saving and economical method to modify existing DNA fragments.

At Entelechon, PCR is routinely integrated with other techniques such as gene synthesis and library construction. Please inquire for a detailed price quote and analysis for your project.

Standard PCR (500 nt)

BIO-003-001

PCR of further 500 nt

BIO-003-002



perfectGene

How to order

If you are interested in Entelechon's gene synthesis services, the next step is easy:

Call our technical hotline

+49 (9405) 96 999 10

or send an inquiry to

perfectGene@entelechon.com

Or visit www.entelechon.com where you will find additional information and be able to order online.



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com





perfectBiology

Oligonucleotide synthesis
Site-directed mutagenesis
DNA libraries
PCR
Plasmid upscaling service
Subcloning

The Popcorn Polymer

Robert Letsinger: Father of Oligonucleotide Synthesis

1963 – 1965, Northwestern University, Evanston, Illinois:

Once the double helical structure of the DNA had been elucidated and the genetic code had been broken, many biologists directed their attention towards synthesizing the newly discovered macromolecules, proteins, and nucleic acids.

Because their chemistry was simpler, the synthesis of proteins developed faster, and in the early 1960s, the chemist Robert Letsinger almost succeeded. But science can be cruel: Robert Bruce Merrifield published his results on the polypeptide synthesis on a polymeric support a few months before Letsinger; he was awarded the Nobel prize in 1984.

This prompted Letsinger, the “father of oligonucleotide synthesis”, to change his research subject – for the benefit of bioresearch everywhere.

He made use of the same approach introduced before for peptide synthesis, attaching one end of the growing chain to a solid support. Although the method worked in principle by 1965, it took a good deal of optimization. One of the problems in the beginning was the swelling of the popcorn polymer, used as support, in some solvents.

This and many other technical problems were solved mainly by Marvin Caruthers from the University of Colorado in Boulder, whose teamwork with Leroy Hood of CalTech led, in the early 1980s, to the commercialisation of the first DNA synthesizers by Applied Biosystems, Inc. (ABI).

Today, synthesis of nucleic acid has come a long way: Billions of oligonucleotides and synthetic genes are produced each year in a tightly controlled and efficient industrial process.

Food Forensics: PCR puts Basmati rice on trial

As Basmati rice, virgin olive oil, and tuna arrive in the docks, a daily concern is the question whether they are the genuine product or maliciously mislabelled imitations?

In 1985, the British geneticist Sir Alec Jeffreys invented a new technique called “genetic fingerprinting”, nowadays a widely applied method in forensics for law enforcement. But it’s not just criminals which possess that individual, clearly identifiable DNA profile.

DNA-based techniques are also used for determining the authenticity of food products: “Original Basmati rice” means produced in the Himalayan foothills and nowhere else. The label “original” on a package of long grain rice that superficially looks like Basmati but misses all its characteristic properties would mean fraudulent misdescription.

Where many chemical techniques fail to identify such fraud, genetic fingerprinting does the trick. Minimal variations in 15 different gene loci, called SSLPs (= small sequence

length polymorphism), can detect genuine Basmati rice. Analysis of SSLPs or SNPs (=single nucleotide polymorphism) are also applied to identify closely related members of the tuna family or different cultivars of olives.

Using DNA fingerprinting, food inspectors can have a closer look at processed meat: Since the BSE crisis, testing for neuronal tissue in processed meat has become a routine procedure. Using sophisticated quantitative PCR methods, scientists succeed in verifying a contamination of processed meat with as little as 0.01% of spinal cord, even if both tissues are derived from the same animal with exactly the same genome. But muscle and non-muscle tissue show differences in the methylation state of their gene promoters - minimal modifications, but sufficient for the detection by PCR.

This is excellent news for the consumer: More and more loopholes in consumer product fraud are closed, and to no small part due to advances in molecular biology.

Oligonucleotide synthesis

Entelechon's success in gene synthesis is owed to no small part to our high quality oligonucleotides. Whereas applications such as sequencing and PCR are quite forgiving when it comes to oligo quality, gene synthesis requires uncompromised oligonucleotide quality. Mismatches and n-1 products prove fatal for the assembly of genes.

Therefore, both the oligonucleotides used in-house as well as those provided to clients undergo rigorous quality control and are synthesized under a tightly controlled process, using fresh chemicals at all times.

If you need randomized oligonucleotides, we can provide those as well. For optimum distribution of nucleotides at the randomized positions, we pre-mix the nucleotides offline rather than let the oligosynthesizer inject the nucleotides simultaneously. We do of course account for the different coupling efficiencies of the four nucleotide types.

Delivery time: Within 2-3 work days

- PCR primers
- Standard oligonucleotides in gene synthesis quality with an extremely low rate of mismatches and n-1 products
- Fluorescent labelled probes (also with multiple labels)
- Oligonucleotides with randomized nucleotide positions

	Scale	0.04 μmol	0.2 μmol	1 μmol
Desalted	Yield OD ₂₆₀	5	16	80
	Product no.	OG-001	OG-002	OG-003
HPLC	Yield OD ₂₆₀	3	10	30
	Product no.	OG-004	OG-005	OG-006

Oligo modifications

We can provide standard oligonucleotides, but also a wide range of chemical modifications and fluorescent labels. If you require a non-standard modification that is not listed here, please inquire – it is very likely that we can provide what you need at very competitive prices.

Examples of 5' Modifications:
Fam, Hex, Biotin

5' modification

OG-007

Examples of 3' Modifications:
Amino, Phospho, Biotin

3' modification

OG-008

Customized primer design

Our bioinformatics unit designs customized oligonucleotides for PCRs, as hybridization probes or for microarrays, covering anything from specific gene sets to whole genomes. Our oligonucleotide design is based on proprietary software developed in-house, thus giving us full control over the process. We routinely adapt the design process to specific client requirements.

Customized primer design

INF-008

Mutation service

Our mutation service allows to introduce one or more site-specific mutations into an existing DNA fragment, including insertions and deletions.

This is useful for a wide range of applications, such as functionality studies, interaction studies, knock-outs, and splice-site studies. Sometimes adapting an existing gene by site-directed mutagenesis is more efficient and faster than complete de novo synthesis.

Site-directed mutagenesis

BIO-001

DNA libraries

For the construction of gene libraries, we can introduce randomized nucleotide positions during gene synthesis.

You are completely free to choose any mixture of nucleotides at any given position, including a uniform distribution of all four nucleotides or asymmetric mixes of two, three or four nucleotides.

Please note that DNA synthesis is not entirely error-free. The end product will contain small amounts of mismatches at the non-randomized positions. The exact ratio of erroneous versus correct products depends on a number of factors such as gene length and sequence composition and cannot be predicted a priori. However, Entelechon's strict quality control and tightly controlled synthesis process ensure the highest possible quality.

All libraries can be delivered as non-ligated PCR products or cloned into a standard or custom vector.

DNA library

BIO-002



perfectBiology

PCR

Although PCR is a stable technique in most molecular biology labs, it still has its pitfalls. Depending on the desired application, PCR setup and development of a robust PCR protocol can be tricky.

Entelechon's advanced molecular biology lab can develop virtually any PCR application. Our long track record of successful implementation of PCR-based protocols helps us to attack very challenging problems, such as the development of multiplex PCR assays that simultaneously and semiquantitatively detect 24 targets.

Although by today's standards a very traditional method, PCR can still be the most time-saving and economical method to modify existing DNA fragments.

At Entelechon, PCR is routinely integrated with other techniques such as gene synthesis and library construction. Please inquire for a detailed price quote and analysis for your project.

Standard PCR (500 nt)	BIO-003-001
-----------------------	-------------

PCR of further 500 nt	BIO-003-002
-----------------------	-------------

Plasmid Upscaling Service

Based on a plasmid provided by the client or made by Entelechon, plasmid production can be upscaled into the milligrams range. Whether you need high amounts of DNA for transfection, DNA vaccination, virus or antibody production or preclinical animal studies, we can provide the required amount and quality under a certified production process.

Standard quality grade

The standard production of plasmid DNA includes:

- Transformation of an E. coli production strain with the target plasmid
- Reproducible quality of plasmid DNA by establishing an individual cell bank
- Cultivation of cells in media without animal-derived compounds
- DNA production by standardized manufacturing processes, including LPS endotoxin removal (< 0.1 E.U. / mg DNA)
- Multi-parameter quality control based on seven independent assays
- Concentration from 0.5 to 5.0 mg/ml

Premium quality grade

In addition to the features of the standard quality grade, this includes:

- Significant reduction of the open circular and linear plasmid forms
- Significant reduction of bacterial chromosomal DNA
- Elimination of RNA contamination by RNase

	Standard grade	Premium grade
5 mg	BIO-004-005	BIO-005-005
10 mg	BIO-004-010	BIO-005-010
20 mg	BIO-004-020	BIO-005-020
50 mg	BIO-004-050	BIO-005-050
100 m	BIO-004-100	BIO-005-100

Vector logistics

On request we store aliquots of your plasmid preparation according to ICH guidelines under GMP conditions and deliver requested amounts on dry ice by overnight courier service.

Additional quality control options

The following QC tests are offered optionally in addition to the standard QC. Depending on the plasmid the specifications given below may change.

- Endotoxin (LPS) < 0.1 E.U. / μg DNA. LAL test (Limulus Amebocyte Lysate)
- DNA homogeneity (ccc content) \geq 90% ccc. Capillary gel electrophoresis
- Bacterial chromosomal DNA. Quantitative PCR
- Proteincontent. BCA test (Bicinchoninic Acid)
- Purity (microorganism). Bioburden
- Identity sequencing (double-strand)

All QC tests are performed according to documented procedures . Optionally, these test can be performed under GLP/GMP conditions.

Endotoxin	BIO-006-001
DNA homogeneity	BIO-006-002
Bacterial chromosomal DNA	BIO-006-003
BCA test	BIO-006-004
Bioburden	BIO-006-005
Identity sequencing	BIO-006-006

Subcloning

We can clone fragments provided by the client or produced at Entelechon into a range of standard target vectors as well as into custom vectors.

Examples of available vectors are:

Cloning vectors	Expression vectors
pSC-A	pET-vectors
pPCR-Script-Amp	pGEX-4T-1, pGEX-6p-1
pCR4-TOPO	pEGFP-C1
pUC18, pUC19	pIRES-vectors
pBluescriptII	pCDNA3.1
	pMAL-c2

Please inquire for other vector options and for subcloning into a vector provided by you.

Subcloning in-house vector BIO-007-001

Subcloning customer vector BIO-007-002



perfectBiology

How to order

If you are interested in Entelechon's molecular biology services, the next step is easy:

Call our technical hotline

+49 (9405) 96 999 10

or send an inquiry to

perfectBiology@entelechon.com

Or visit www.entelechon.com where you will find additional information.



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com





perfectProtein

Protein expression
Screening of expression clones
Protein characterization
Protein purification
Quantitative proteomics
Synthetic peptides



perfectProtein

The Importance of Working with Insulin

Banting & Macleod, Sanger, Yalow: Three Nobel Prizes for Insulin

December 1923, 1958 and 1977, Stockholm, Sweden:

The Nobel Prize is awarded, and each of these years the molecule involved is the same: Insulin, the hormone that regulates the concentration of sugar in our blood. In 1923, Frederick Grant Banting and John James Richard Macleod are awarded the prize for the discovery of insulin. Indeed, much work on the subject has been done before the two. Together with Charles Best and James Collip, they succeed in purifying insulin from bovine pancreas and are able to help a young diabetic of fourteen years, without the often deleterious side-effects of earlier products. Perhaps by sheer luck August Krogh, the winner of the Nobel Prize in Physiology and Medicine in 1920, meets Macleod in Toronto and receives from him the protocol for insulin purification. Using the protocol, he is able to help his diabetic wife.

Not surprisingly, he nominates Macleod for the Nobel award – which he wins only one year after the experiments. Krogh also starts an endeavour in 1923, which should later become Novo Nordisk, world leader in diabetes care.

Since 1923, insulin is commercially available and in 1926 it is one of the first proteins to be crystallized in pure form. In 1958, the fact that it was the first protein to be fully sequenced is the reason for Frederick Sanger's first Nobel Prize (he receives a second one in 1980). In 1963, insulin is the first protein to be chemically synthesized. Rosalyn Sussman Yalow, in the mid-1950s, develops a method for the radio-immunological determination of insulin in the blood – which wins her the Nobel Prize in 1977.

Insulin not only brought its researchers fame, but also economic success: it is the first protein to be manufactured biotechnologically in 1978, constituting the cornerstone of Genentech's success.



perfectProtein

Mottai-nai!

Koichi Tanaka: Soft Desorption Ionisation Method for Mass Spectrometry

February 1985, Shimadzu Corporation, Kyoto, Japan:

The young physicist Koichi Tanaka realizes he has made a mistake: instead of using acetone for the dilution of Cobalt Ultra Fine Metal Powder, he took the wrong bottle and mixed the UFMP with glycerol. But as his grandmother used to say, “Mottai-nai” - “What a waste!”, he does not throw the mixture away, but puts it into the mass spectrometer nonetheless. This mistake is the reason why Tanaka, seventeen years later, should be the first person ever without a post-bachelor’s degree to win a Nobel Prize in a scientific field.

Mass Spectrometry had for a long time been used in the analysis of peptides, but because of the difficulty to ionise a substance without fragmentation to small pieces, it could not be used for the analysis of macromolecules like proteins.

Nevertheless, there was of course demand for the MS-based characterization of proteins. Thus Shimadzu’s Director decided to let a group of five researchers work on this problem. The team indeed made great advances in the optimization of the mass spectrometry parameters, but they failed to develop a matrix allowing the ionisation of an analyte without destruction of the structure – until Tanaka’s blunder.



perfectProtein

Protein expression

Entelechon provides protein expression in a range of versatile expression systems. Our expertise in protein production as well as solid protocols for expression and purification of even very difficult proteins make the process fast and efficient.

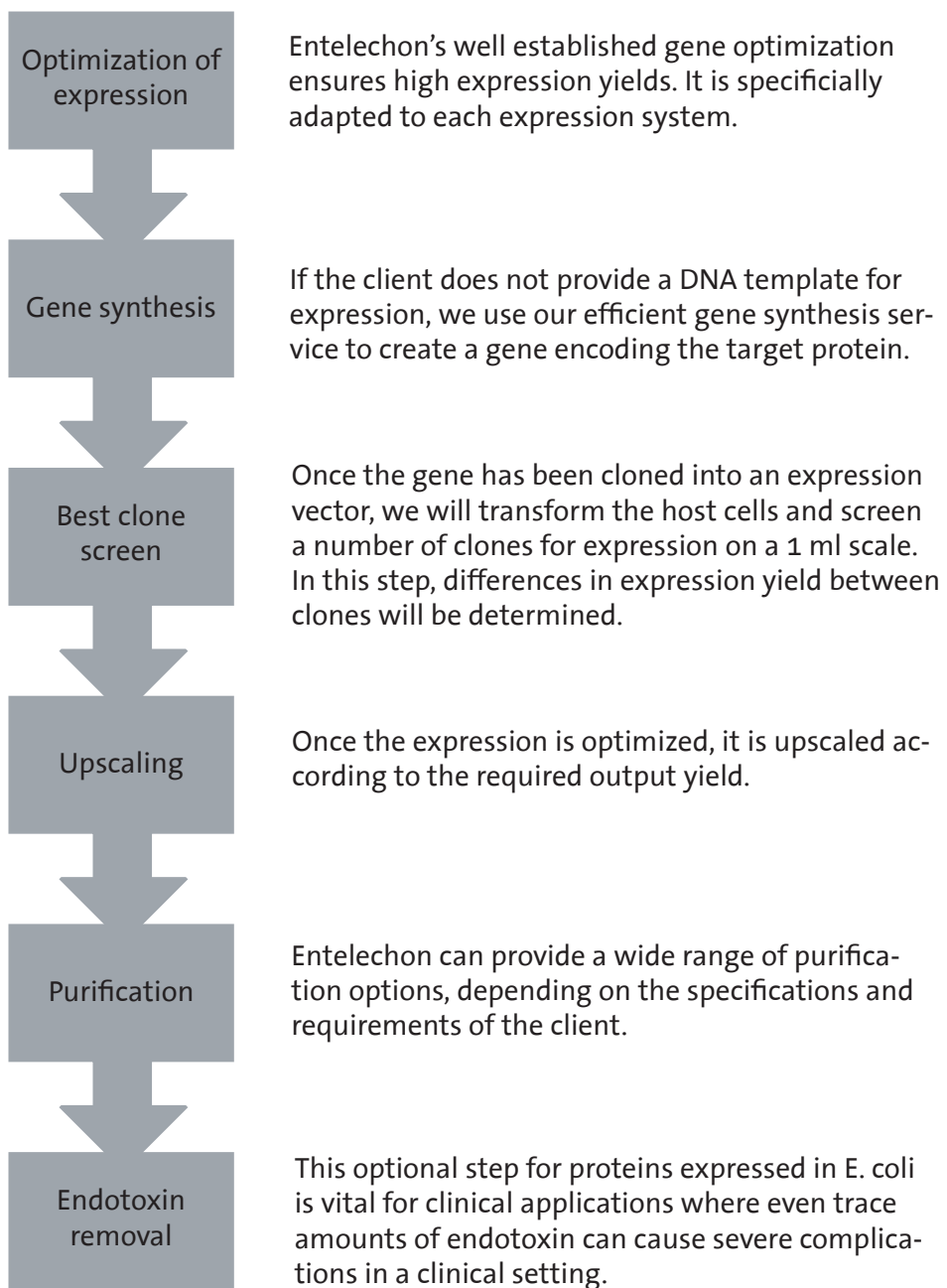
We have a track record of handling proteins with unusual properties, such as extremely hydrophobic proteins, structural proteins with a strong tendency towards aggregation or membrane proteins.

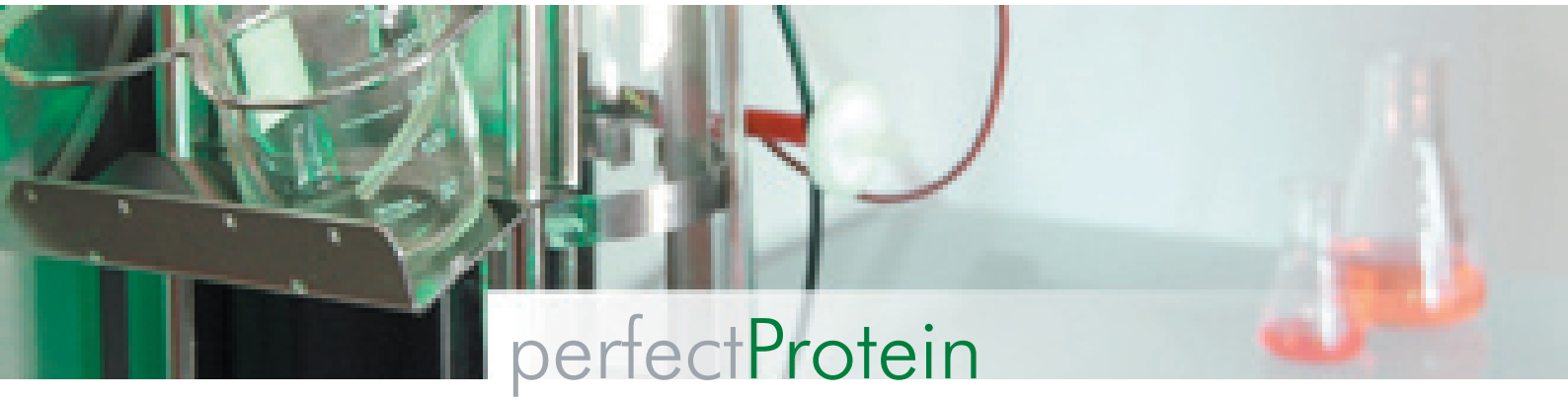
Expression can easily be scaled from 1 ml to 300 litres batches. Gene expression follows a milestone plan, with complete flexibility for the client: At each milestone, the client will receive a detailed report, including clear, and precise recommendations on the next step. The decision of whether and how to continue is always in the hands of the client.

In addition to the complete workflow we can also provide single milestones separately. The milestones are described on the following page.



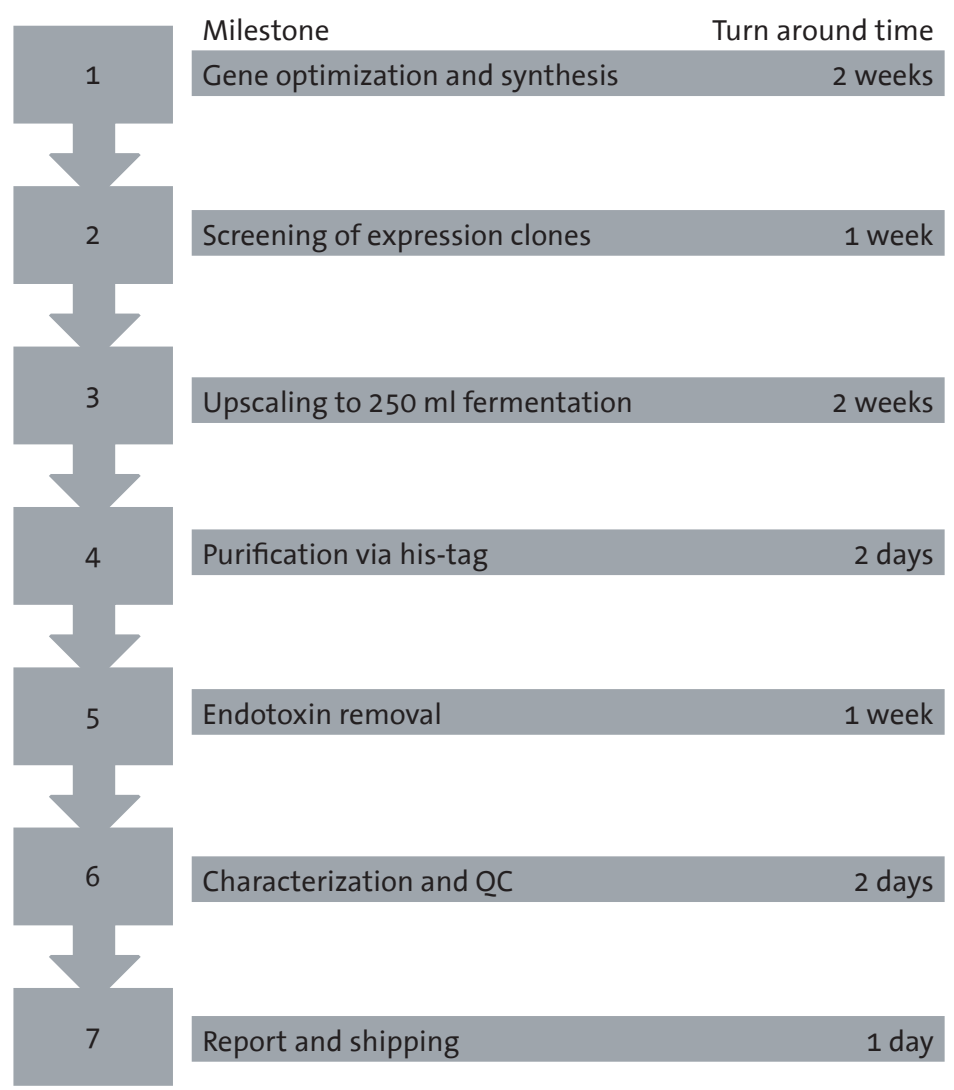
perfectProtein





perfectProtein

Here is a typical workflow for the expression and purification via his-tag of a non-native protein in E. coli:





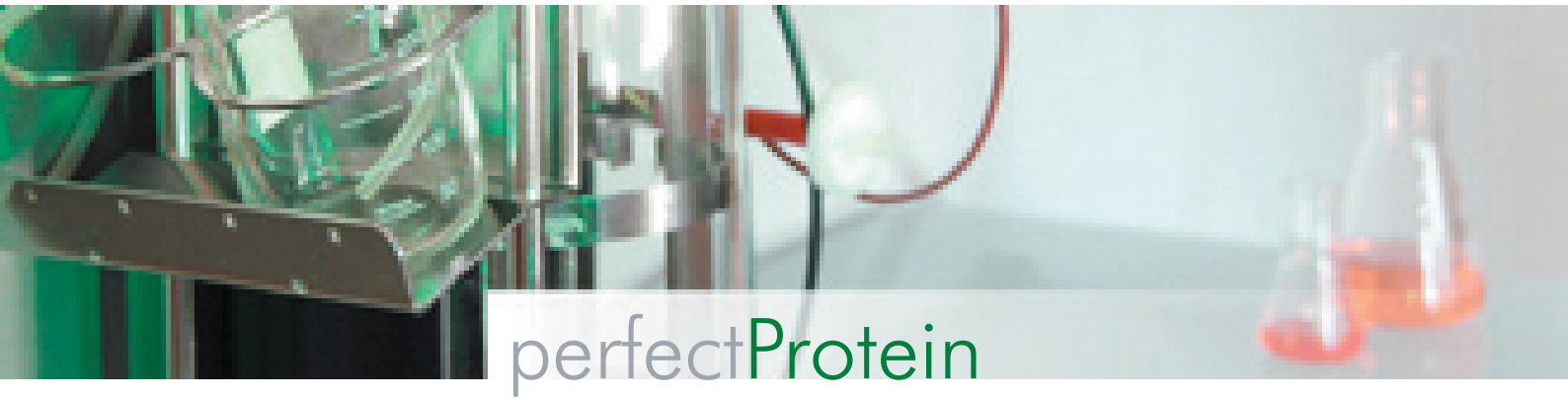
perfectProtein

Entelechon maintains a versatile collection of expression systems, tailored towards the particular requirements of each individual client. The Entelechon team is happy to provide detailed advise on the advantages of each expression system and will help you in deciding which one best suits your project:

- Escherichia coli
- Bacillus subtilis
- Saccharomyces cerevisiae
- Schizosaccharomyces pombe
- Pichia pastoris
- Insect cell lines
- Mammalian cell lines

We determine the expression system best suited for a particular project in close discussion with the client. Depending on the scope of the project, it may be warranted to start with a simple expression system such as E. coli for an initial proof of concept and to later move to an expression system with enhanced features, such as mammalian cell lines. Available expression systems are:

E. coli expression	PRO-001-001
B. subtilis expression	PRO-001-002
Yeast expression	PRO-001-003
Insect cell line expression	PRO-001-004
Mammalian expression	PRO-001-005



perfectProtein

Screening of expression clones

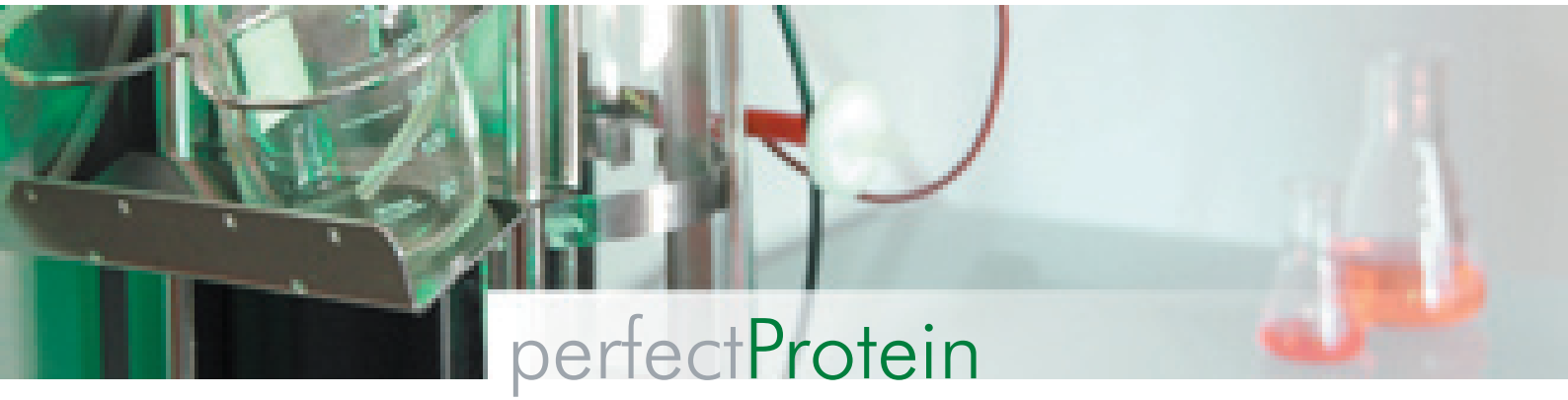
Given a set of recombinant clones of an expression host, Entelechon can perform the screening of these clones for expression yield. This screening is the starting point of upscaled gene expression, but can also be part of the post-processing of gene synthesis, so that the client gets a proof of the actual expression level of a newly synthesized gene.

In the screening, a set of at least seven clones will be selected and subjected to gene expression on 1 ml scale. Expression will be detected by SDS-PAGE. Optionally, with our extended service, vital parameters of the expression protocol - such as variations in temperature, different strains, different plasmids or the induction regiment - can be applied.

The client receives the best expressing clone along with a detailed report of the expression conditions and resulting yields. The expression screen provides a good indication of how to proceed with upscaling and what scale of expression is required for the targeted protein quantity.

After having screened for the best expressing clone, Entelechon can also provide a solubility study for the target protein.

Standard best clone screen	PRO-002-001
Extended best clone screen	PRO-002-002
Solubility study	PRO-002-003



perfectProtein

Protein purification

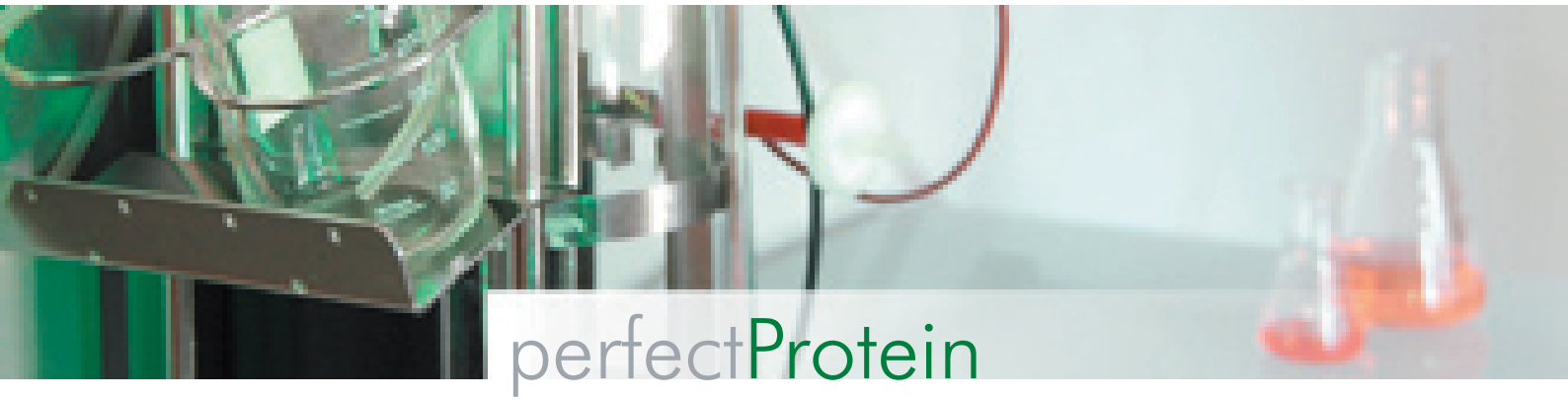
This is the most versatile and most demanding step of the expression process. Entelechon has developed a number of protocols that give us plenty of options for the purification of virtually any protein. Starting from simple ion exchange chromatography via affinity chromatography up to complex strategies based on stacking several purification steps and using multiple affinity tags, our purification procedures are adjusted to the desired purity and nature of the target protein.

Using rigorous QC based on chromatography, SDS-PAGE, Western Blot, ELISA, and mass spectrometry, not only can we ensure the required level of purity, but prove it.

If a protein is required in its native, functional state, we leverage our expertise in reconstituting even delicate protein structures. We like to work closely with the client in developing suitable assays for functionality, often incorporating and adapting pre-existing tests developed by the client.

Available purification options are:

Affinity chromatography	PRO-003-001
Ion exchange chromatography	PRO-003-002
Gel filtration	PRO-003-003
hic - hydrophobic interaction	PRO-003-004
Refolding of denatured protein	PRO-003-005
Sample concentration	PRO-003-006
Protease digestion	PRO-003-007
Endotoxin removal	PRO-003-008
Lyophilization	PRO-003-009



perfectProtein

Protein characterization

Entelechon offers several services for protein characterization. Besides standard QC methods such as SDS PAGE, Western Blot and UV/Vis spectrometry, total amino acid analysis provides an accurate quantification of a pure protein, whereas N-terminal sequencing allows to quickly confirm the identity of a protein or peptide.

At least 10 pmol of pure protein/peptide with >80% purity are required, to be provided on PVDF membrane or in a low salt buffer solution, free of amines.

Sequencing will be performed stepwise, starting with the first five amino acids. We follow the Edman degradation and can sequence up to 10 amino acids.

If sequencing of the initial five amino acids is unsuccessful, a setup fee applies.

For the total amino acid analysis 1 - 5 µg of the protein, 0.2 - 1 µg of the peptide, respectively, are required.

Please inquire about the specifications for SDS PAGE, Western Blot and spectrometry.

SDS PAGE	PRO-004-001
Western Blot	PRO-004-002
UV/Vis spectrometry	PRO-004-003
N-terminal sequencing	PRO-004-004
Total amino acid analysis	PRO-004-005

Quantitative proteomics

Entelechon's subsidiary PolyQuant provides the multiplex quantification of proteins based on its proteomics platform QconCAT™. This technology allows for the simultaneous and accurate quantification of a set of up to 100 proteins, originating from arbitrary matrices such as blood serum, cell extracts or soil samples.

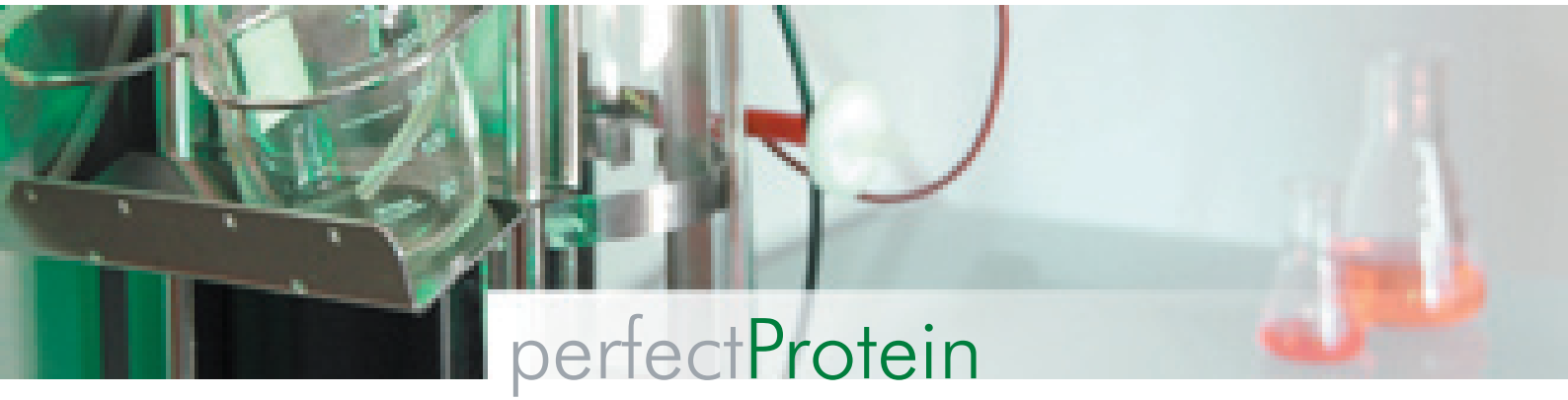
QconCAT™ provides absolute protein quantities that can be easily compared across independent experiments and between multiple labs. Assay development is very fast with a typical turn around time of 12 weeks for a *de novo* assay. This is a significant advantage over ELISA, where assay development can take many months.

PolyQuant offers protein quantification as a fee-for-service, but also provides quantification reference standards for in-house use as well as off-the-shelf quantification kits. If you want to learn more, please contact

PolyQuant GmbH
www.polyquant.com
info@polyquant.com
Tel. +49 (9405) 96 999 10

or send an inquiry to Entelechon.

The QconCAT™ technology can be perfectly combined with the gene synthesis, protein expression, and peptide synthesis services from Entelechon. It can be seamlessly integrated into Entelechon's project workflows and clients can interact with Entelechon as the integration point of both PolyQuant's and Entelechon's own services.



Synthetic peptides

Entelechon provides chemically synthesized peptides for a number of applications, such as antigen-antibody interaction studies, epitope mapping, receptor studies or enzymatic assays. Peptides come at different quantities and purity levels - each perfectly adjusted to the specific requirements of your project.

A number of covalent modifications and non-standard amino acids can be incorporated into the peptides.

The standard length is 8 to 30 amino acids. Longer peptides are readily available at a length-dependent extra fee.

Purity ranges from crude synthesis preparation to >95%. Quantities range from 2 mg to 1000 mg.

	crude	> 70 %	> 80 %	> 90 %	> 95 %
2 mg	PEP-001-001	PEP-001-002	PEP-001-003	PEP-001-004	PEP-001-005
5 mg	PEP-001-006	PEP-001-007	PEP-001-008	PEP-001-009	PEP-001-010
10 mg	PEP-001-011	PEP-001-012	PEP-001-013	PEP-001-014	PEP-001-015
20 mg	PEP-001-016	PEP-001-017	PEP-001-018	PEP-001-019	PEP-001-020
50 mg	PEP-001-021	PEP-001-022	PEP-001-023	PEP-001-024	PEP-001-025
100 mg	PEP-001-026	PEP-001-027	PEP-001-028	PEP-001-029	PEP-001-030
200 mg	PEP-001-031	PEP-001-032	PEP-001-033	PEP-001-034	PEP-001-035
500 mg	PEP-001-036	PEP-001-037	PEP-001-038	PEP-001-039	PEP-001-040
1 g	PEP-001-041	PEP-001-042	PEP-001-043	PEP-001-044	PEP-001-045



perfectProtein

Modifications, for instance:

- Acetylation (N-terminal)
- Amidation (C-terminal)
- Benzyloxycarbonylation (CBZ)
- Fatty acid (N-terminal)
- Formylation (N-terminal)
- p-Nitroanilide (pNA)
- Succinylation (Suc)(N-terminal)

Please inquire for other modifications.

Modifications

PEP-002

Special amino acids, for instance:

- D-Ile, D-Orn, D-Arg, D-Cys, D-Gln, D-Ser
- Nva, Nle, Hse, Hcy, Pen, pGlu, Orn
- Dinitrobenzoylation (Lys)
- Phosphorylation (Tyr, Ser, Thr)

Please inquire for other special amino acids.

Special amino acids

PEP-003

Conjugations:

- KLH coupling
- BSA

Coupling

PEP-004



perfectProtein

Fluorescence/dye labeling:

- Aminocoumarin (N-Terminal)
- Biotin (N-Terminal)
- DTPA (N-Terminal)
- HYNIC (N-Terminal)
- FITC (N-Terminal)

Quenched fluorescent peptides:

- Abz/ Tyr (3-NO₂)
- EDANS/ DABCYL

Flourescence/dye labeling	PEP-005
---------------------------	---------

In addition to standard peptides, we offer cyclic peptides and multiple antigenic peptide (MAP) systems. MAPs are asymmetrical branched structures of multiple peptides which strongly increase the antigenicity of a given peptide epitope.

4 Asymmetric branches	PEP-006-001
8 Asymmetric branches	PEP-006-002
Disulfide bridged	PEP-007-001
Amide cyclic	PEP-007-002



perfectProtein

How to order

If you are interested in Entelechon's protein expression, protein analysis, and peptide synthesis services, the next step is easy:

Call our technical hotline

+49 (9405) 96 999 10

or send an inquiry to

perfectProtein@entelechon.com

Or visit www.entelechon.com where you will find additional information and be able to order online.



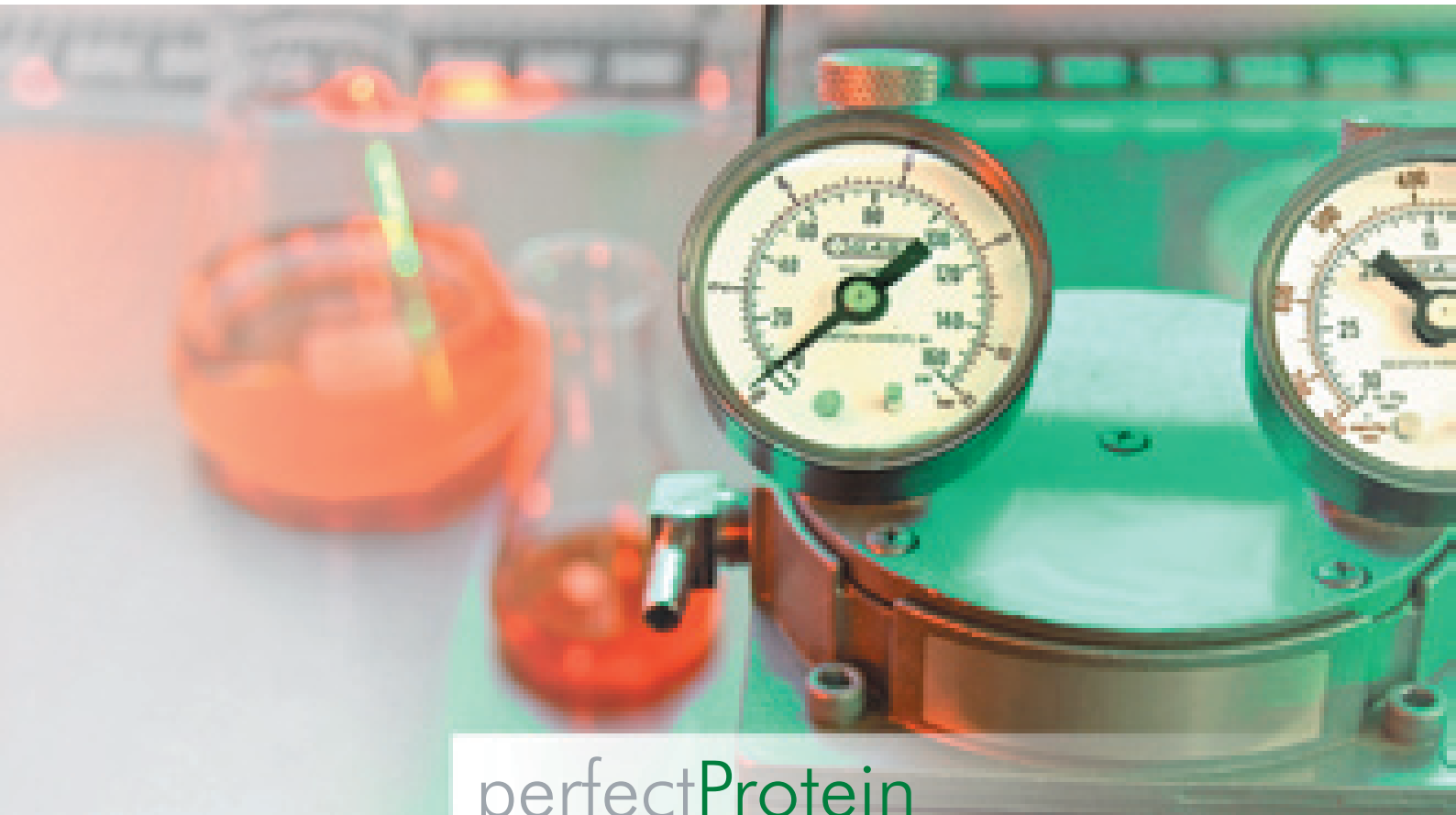
Contact data

Entelechon GmbH
St. Veit-Weg 2
93053 Regensburg
Germany

Tel. +49 941 6981810
Fax +49 941 6981828

contact@entelechon.com

www.entelechon.com



perfectProtein

Antibodies

Milstein and Köhler: The hybridoma technology

December 1974, Cambridge: Georges J.F. Köhler, a postdoctoral fellow of César Milstein, had given eternal life to normal antibody producing B cells: He realized that it should be possible to convey immortality to B lymphocytes by fusing them with ever-proliferating myeloma tumor cells.

Several selection and purification steps later Köhler held in his hands a truly remarkable specimen: Immortal hybridoma cells that produce exactly one type of antibody: The first monoclonal antibody (MAb).

In 1975 they published their hybridoma technology in "Nature" and in 1984 Milstein and Köhler shared the Nobel Prize for Physiology or Medicine with Niels K. Jerne. They produced the first monoclonal antibody against the T lymphocyte subset antigen CD4 and, in co-

operation with L. Herzenberg who had just introduced FACS analysis to biology, they soon realized the synergy between FACS and MAbs. Milstein & Co produced their MAbs in hybridoma cells.

Today, also transgenic mice are used besides hybridoma cells as a source for MAbs. Both techniques are expensive and provide only limited quantities of monoclonal antibodies.

This has prompted a number of biotech companies to attempt production of MAbs in plants: After introducing human antibody genes into tobacco plants, these genetically altered crops are able to produce fully functional human MAbs – fittingly dubbed "plantibodies".



perfectProtein

I say –ximab, you say –zumab! Let’s say nano?

To the layman, the names of therapeutic antibodies sound more like Aztec deities than like a scientific nomenclature.

Rituximab and Bevacizumab: These tongue-twisters are names of therapeutic, monoclonal antibodies (MAb), one for treating non-Hodgkin lymphoma and the second one is the first commercially available angiogenesis inhibitor. Manufacturers of MAb-based therapeutics are free to choose the first part of the name only: Ri- or Beva-. The next syllable describes the target of the antibody: “tu” means tumor and “ci” cardiovascular. The suffix “mab” is certainly easy to understand, but the one or two preceding letters may still be mysterious: They indicate the origin of the antibody. The “o” in Tositumomab, also a treatment for lymphoma, conjugated with a radioactive isotope, denotes production in mice.

But our immune system recognizes non-human antibodies and so they may be cleared off rapidly. In case of an application with radioactively labelled MAbs this is surely a wanted reaction, but foreign antibodies also provoke unwanted allergic reactions. Therefore scientists engineered MAbs more

immunotolerable - by fusing the murine variable regions to human constant regions. So Rituximab is a mouse-human chimeric antibody – marked by the “xi” in its name.

In contrast the infix “zu” identifies Bavacizumab as a humanized antibody. This means that not only its constant, but also parts of the variable regions are of “human” origin. Disadvantages of traditional therapeutic MAbs are their need for low storage temperatures and their size: They are too big to enter solid tumors or to easily permeate the blood-brain barrier – the latter would be the premise for any therapeutic application targeting the central nervous system, such as in Alzheimer’s Disease.

So-called nanobodies, produced in dromedaries, consist solely of a pair of heavy chains. Despite having only about a tenth the size of “normal” antibodies, they possess a virtually equal affinity to their targets. Because nanobodies show more resistance to pH and heat, they may be suitable for oral administration. So far, the dromedary is not a common source for antibodies, but this is surely bound to change. Perhaps we should start to look for a syllable denoting “humanized dromedary nanobodies”.

Polyclonal antibodies

Entelechon's antibody production service can provide antibodies from rabbit, mouse, and chicken egg yolk. We can work from peptides or proteins synthesized in-house or can use material provided by the client. These services can be complemented by affinity purification. Quality control consists of a determination of ELISA titers.

Rabbit antiserum

This is the most common antibody source. For one rabbit, either 0.2 - 0.6 mg of protein or 1 - 2 mg of peptide are required. From one rabbit, 35 - 80 ml serum can be collected. The delivery time for protein antiserum is 30 - 55 work days, for peptide antiserum it is up to 90 work days.

Antiserum from one rabbit

AB-001

Mouse antiserum

When only small amounts are required, murine antiserum is the best choice. From one mouse 0.5 - 1 ml antiserum will be obtained.

Antiserum from one mouse

AB-002

Chicken egg yolk antiserum

If you require antibodies directed against highly conserved mammalian proteins, chicken would be the best choice. It is a cost effective method if you need high yields of antibodies from limited amounts of available antigen. For one hen 0.2 - 0.6 mg protein or 1 - 2 mg peptide are required. From 10 immune eggs 200 - 1000 mg IgG will be collected (> 90% purified). First egg collection starts around day 55-65 after immunization and can last up to 600 days.

Chicken egg yolk antibody

AB-003

Preimmune serum can be collected upon the customer's requirements.

Additional services

Peptide synthesis for antibody generation

We synthesize peptides required for antibody production. The synthesis will take 3 - 5 weeks. A HPLC chromatogram and MS spectrum will be provided as quality control.

Antibody peptide synthesis

AB-004

Monospecific antibodies

Enrichment of antibodies by affinity purification in order to increase the specificity. Delivery time will be 5-10 work days.

Affinity purification

AB-005

Sample requirements

For **proteins**, we require about 400 µg of the protein with a concentration of 0.2 mg/ml or higher.

Peptides must be conjugated to a large carrier molecule, e.g. KLH, BSA, ovalbumin or polydextran, in order to make them immunogenic.



perfectProtein

Antibodies and quantitative proteomics

Entelechon's subsidiary PolyQuant offers an innovative method for the quantification of proteins, called QconCAT™. By spiking an isotope labeled polypeptide into the analyte solution and performing a tryptic digest, proteotypic peptides are released from the target proteins as well as from the heavy reference polypeptide.

This technique is in itself very useful for the precise quantification of proteins from arbitrary matrices.

However, in combination with antibody-based affinity enrichment, the sensitivity of QconCAT™ can be significantly enhanced. If you are interested in combining Entelechon's antibody production and PolyQuant's QconCAT™ technology, we are happy to provide detailed information. Please contact us at the coordinates on page AB-7, or at PolyQuant:

PolyQuant GmbH
www.polyquant.com
info@polyquant.com
Tel. +49 (9405) 96 999 10



perfectProtein

How to order

If you are interested in Entelechon's antibody services, the next step is easy:

Call our technical hotline

+49 (9405) 96 999 10

or send an inquiry to

perfectProtein@entelechon.com

Or visit www.entelechon.com where you will find additional information and be able to order online.



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com





perfectSilico

Datamining

Database setup

Software development

Microarray oligonucleotide design

Custom bioinformatics solutions

Clustering and homology analysis



perfectSilico

The Scientist, the Tree, and the Squirrel

Marshall Nirenberg: Discovery of the Genetic Code

Saturday, May 27, 1961, three o'clock in the morning – National Institutes (NIH) of Health, Bethesda, Maryland:

After the double-helical structure of the DNA had finally been elucidated by Watson and Crick, one big mystery remained: how is the genetic information translated into living structures?

In the 1950s, the scientific community was involved in a race for the solution to this conundrum – everybody wanted to be the first to break the code of life. In Cambridge, the RNA Tie Club was founded by George Gamow; this hand-picked group of twenty scientists – one for each amino acid, and every one wearing a corresponding tie - met to discuss their ideas.

At the New York University Medical School, a large research team around the Nobel laureate Severo Ochoa was working on the coding problem. And at the NIH in Bethesda, Maryland, Marshall Nirenberg and Heinrich Matthaei were conducting experiments to show that RNA could trigger protein synthesis.

Late one night in May 1961, they succeeded: when the scientists put synthetic RNA consisting of uracil only, to a cell extract from E. coli, the cell-free system produced polypehenylalanine. After years of hard work, the first codon was finally known!

But still, the other codons had to be identified as well. To help Nirenberg in this demanding task, NIH joined all its forces: scientists put aside their own work to go into Nirenberg's lab and assist him in deciphering the code. In 1965, the work was completed, and in 1968 Nirenberg, as the first intramural scientist of the NIH, won the Nobel Prize.

One question that occupied the scientist afterwards was whether the code was universal or specific to bacteria only – or even to the particular species they had worked with. As Nierenberg found out, the code is indeed universal. He later recalled looking out the window of his office, pondering the astounding fact that he, the tree, and the squirrel are so similar on a biological level...

Codon Usage: Wobble and Hobble

You think the genetic code is universal? Well, it is in general but not entirely...

The genetic code is degenerated: most of the amino acids are encoded by more than one base triplet; different species tend to prefer different subsets of these codons, differing in the last base of the triplet.

Mutations in these so-called “wobble” bases were originally thought to be evolutionarily neutral. However, further research found that this codon usage does indeed matter. For instance, DNA vaccination against HIV is more efficient if the codon usage is adapted to that of highly expressed mammalian genes. And recent results show that even the structure of a protein can be changed by “silent” mutation: the transcription machinery within a cell may not be optimized

for the rare codons resulting from the mutation. This may lead to a change in the rate of translation, causing a kind of “hobbling” or translational arrest in the process, that can cause alternative protein folding.

The differently folded protein may suffer from a change in its function and can even be the cause of disease. When the declared adversary of the neutral theory of evolution, Ernst Mayr, claimed some decades ago that silent mutations could not be neutral, that these mutations had to have an effect on the organism that selection can “see” even if the coded amino acid does not change, he met fierce opposition and was accused of dogmatism.

It is clear today that he was right all along and that his faith in the principles of selection was justified.



Datamining

Entelechon uses advanced in-house developed software solutions for the efficient datamining of databases. We can work with public databases as well as those provided by the client.

Datamining can result in a detailed statistical report, but also in the creation of a small database tailored towards the client's problem. Results can be provided in text format, as Excel tables, in a number of more specific formats (such as XML, Genbank, etc) or can be fed into existing data warehouse solutions.

Our workflow is very modular and can be adjusted by incorporating new, client-specific software modules. In this way, we can easily include new data sources, proprietary data format or specific documentation requirements.

Datamining

INF-002

Database setup

Existing databases are often quite complex to deal with and in many instances a dataset reduced in scope or depth may accelerate workflows significantly. Entelechon can create specialized databases tailored to your requirements quickly and efficiently.

Using our datamining tool chain, we can produce databases in virtually any format, such as MySQL, PostgreSQL, XML or as flat text files.

Database setup

INF-003



Software development

Entelechon has a track record of over ten years of successful software development in bioinformatics. A growing collection of bioinformatics modules affords us with a powerful tool set to rapidly develop solutions.

We are ready to disclose the full source code of a software application written for a client. Solutions are accompanied with detailed developer and user documentation. We are happy to interface with existing solutions and standards.

Software development at Entelechon is goal-driven and cost-effective. You will be surprised how affordable and accessible the development of customized tools can be.

A software project at Entelechon can be of any scale, from a one-week project to produce a scratch solution to large-scale projects spanning many months and involving a team of developers. No matter how large a project is, we put strong emphasis on a detailed and open communication with the client at all stages. Our team of molecular biologists and software developers works closely with the client to develop the specification and requirements.

Software development

INF-004

Custom bioinformatics solutions

In addition to projects that require software development from scratch, there are many scenarios that are already covered by solutions available at Entelechon. In general, such projects can be performed very quickly and at very moderate costs.

Custom solutions

INF-005



perfectSilico

Microarray oligonucleotide design

Based on an advanced and highly customizable workflow for oligonucleotide design developed in-house, Entelechon can provide oligonucleotide probes for microarrays and other applications.

Microarrays can have widely varying specifications and our software has been designed with this flexibility in mind: We can easily adapt to the client's specifications, be it the avoidance of cross hybridization against custom targets, physical properties of oligonucleotides or the layout and content of the documentation.

We do, of course, also accept oligonucleotide design projects for other purposes, such as PCR primers, multiplex PCR primers or hybridization probes.

Oligonucleotide design

INF-006

Clustering and homology analysis

Today's high throughput analytical methods and next generation sequencing result in vast amounts of sequence data. In order to make sense of this wealth of information it needs to be postprocessed.

Entelechon can offer advanced methods for clustering and for the identification of homologies within a large set of protein or DNA sequence data. At the same time, our clustering solution is flexible and can easily be adapted to your requirements.

You can choose from a range of clustering strategies, methods for the detection of homologies, and from many different options for documentation and output format.

Clustering/homology analysis

INF-007



perfectSilico

How to order

If you are interested in Entelechon's bioinformatics services, the next step is easy:

Call our technical hotline

+49 (9405) 96 999 10

or send an inquiry to

perfectSilico@entelechon.com

Or visit www.entelechon.com where you will find additional information.



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com





Technical Bulletin 1

Randomized DNA libraries



Technical Bulletin 1

Synthesis

The success of a DNA library depends on a well planned design strategy. The aim is to find a good compromise between the desired number of amino acid variations that are to be explored and a sequence that can be synthesized and handled efficiently.

Essentially there are two aspects to a DNA library design: Synthesis of the library and representation of all desired variants. As for synthesis, the randomized positions must be laid out in a way that still allows the assembly of the full-length gene. For instance, clustering too many consecutive randomized positions will hinder synthesis. We recommend not more than eight consecutive randomized nucleotides.

...ATTCGNNNNNNNNCTT... ✓
...ATTCGNNNNNNNNNNCTT... ✗

(Fig. 1: Upper row shows recommended randomized nucleotides, lower row shows not recommended randomized nucleotides. We use the IUPAC codes for wobble nucleotides, see Appendix.)

Even then, randomized nucleotides should be separated by at least 20 nucleotides. You can organize randomized nucleotides in blocks.

...ATT **NNANNCN** TTGATTGCCTACCAGTTAGCGGAT **NNTNNCNN** ATTG...

(Fig. 2: Minimum separation of randomized nucleotides in blocks.)

A common strategy is to randomize the first two positions of a codon and leave the third one constant. However, please note that interspersing fixed nucleotides at every third position does not make synthesis much easier, therefore such a stretch of NNX codons is considered a consecutive block of randomized nucleotides (see again Fig. 2).



Technical Bulletin 1

In most cases, the aim is to represent a set of amino acids at a given position. Please note that for most sets of amino acids, it is impossible to avoid nucleotide combinations that encode additional amino acids or even stop codons. For example, to represent all 20 amino acids, the best combination is NNK or NNS, which unfortunately codes for the amber stop codon TAG as well.

Using a subset of the four nucleotides, such as G and T (IUPAC code “K”), can reduce the complexity of the library significantly and is therefore advisable where possible. However, please keep in mind that two-nucleotide combinations are almost as difficult in synthesis as four-nucleotide combinations.

The total number of random nucleotide positions in a given sequence is limited both by synthesis and by the screening phase. For synthesis, the limit depends on the exact layout of the randomized positions, but is generally in the range of 10-20 nucleotides. For screening, the limit depends on how well the screening protocol can be automated, whether there is a selection step that works without manual intervention and how many DNA molecules can be synthesized and processed.

The overall length of a randomized DNA sequence should also be limited. Very long sequences are difficult to synthesize even without wobble positions. This difficulty affects the quality of the DNA library. In some cases, very long genes with random positions cannot be correctly assembled at all. A good rule of thumb is to limit randomized genes to 800-1000 bp in length.



Technical Bulletin 1

Non-uniform nucleotide distributions

By default, all nucleotides at a randomized position are represented in equal amounts. However, in some cases it may be desirable to skew that distribution. For instance, when a sequence contains many randomized positions, it could be desirable to reduce the probability of a certain variation at each position, so that the overall probability of the sequence to contain two or more 'mutations' is reasonably low. Therefore, distributions such as 80% A, 20% T or 85% A, 5% C, 5% G, 5% T are often used. The distribution of nucleotides can easily be adjusted accordingly in the oligonucleotide synthesis process.



Technical Bulletin 1

Library screening

When synthesizing a DNA library, Entelechon takes great care to use oligonucleotides of the highest possible quality. We have developed specific synthesis protocols which ensure a low error rate and a uniform distribution of randomized nucleotides.

However, there is a technical limit for the quality of a randomized DNA library. A library will always contain a certain level of mismatches ('mutations') at non-randomized positions, due to the limited perfection of the oligonucleotide synthesis process. Typically this affects less than 20% of all DNA molecules, but this number may be significantly higher or lower for a particular library.

Also, the exact distribution of nucleotides at a particular randomized position usually deviates to some extent from the theoretical distribution. Often the distribution is still within a reasonable margin, such as 20:35:15:30 instead of 25:25:25:25. However, this cannot always be guaranteed and a possible deviation should be compensated by a reasonably high safety factor for the number of contained variants (see below).



Technical Bulletin 1

Screening

The number of variants increases exponentially with the number of randomized positions, and is generally calculated as

$$n = 2^{c_2} * 3^{c_3} * 4^{c_4}$$

where c_2 , c_3 , and c_4 are the number of positions with two, three, and four wobble nucleotides respectively. For example, a sequence with four N positions yields 256 variants, but a sequence with ten N positions already yields over a million variants.

More variants not only make synthesis more difficult, but they also increase the risk that the library does not cover all variants and is therefore skewed. If the number of variants becomes inconveniently large, it is advisable to consider the synthesis of several libraries, each covering a smaller number of randomized positions. This does of course not allow to select variants based on the interaction between separate positions, but it could be used as a first step, in order to identify positions which are likely to have an impact on the overall 'fitness' of a gene or protein.

An important consideration is whether a library can represent all intended variations. In order to verify this, first calculate the number of variants according to the formula above. The randomized DNA follows a binomial distribution, which allows to calculate the number of individual DNA molecules m needed for a given number of variants and a desired probability p to represent each variant at least once in the library:

$$m = \log_{(1-p)/n} ((1-p)/n)$$

p should be close to 1, e.g. 0.9999. Now, we have to ensure that the library contains at least this many DNA molecules. We get the required amount in mols by dividing by the Avogadro constant $6.0221415 * 10^{23}$ mol⁻¹. Dividing the synthesis scale (for instance, the default synthesis



Technical Bulletin 1

scale for genes of up to 1000 bp at Entelechon is approximately 10 nmol) by this amount yields a final factor.

If that factor is greater than one, the library should contain all variants at least once with the previously specified probability. Note that a number of steps in the synthesis and screening process can distort the library distribution. For instance, cloning into a vector, PCR, and propagation through host cells will inevitably lead to positive and negative selection of some variants.

Therefore, the calculated factor should be significantly higher than one. If, for example, we expect the underrepresentation of a particular variant due to negative selection about a factor of ten, the calculated value should be greater than ten. In practice, a factor of 10^4 to 10^6 is recommended.

We assist with the library design

A library design has many aspects, and oftentimes it is not easy to select the perfect design. Entelechon has a long track record of creating libraries successfully, and we are happy to design a library for you or review a design made by you.



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com

entelechon





Appendix

Genetic code
Common restriction enzymes
Spectroscopic extinction coefficients
Centrifugal conversions
DNA and protein gels
Fusion proteins and tags
Protease cleavage sites
Sequencing primers



Appendix

Genetic code

1 \ 2	T	C	A	G	2 \ 3
T	Phe (F)	Ser (S)	Tyr (Y)	Cys (C)	T
T	Phe (F)	Ser (S)	Tyr (Y)	Cys (C)	C
T	Leu (L)	Ser (S)	STOP	STOP	A
T	Leu (L)	Ser (S)	STOP	Trp (W)	G
C	Leu (L)	Pro (P)	HIS (H)	Arg (R)	T
C	Leu (L)	Pro (P)	HIS (H)	Arg (R)	C
C	Leu (L)	Pro (P)	Gln (Q)	Arg (R)	A
C	Leu (L)	Pro (P)	Gln (Q)	Arg (R)	G
A	Ile (I)	Thr (T)	Asn (N)	Ser (S)	T
A	Ile (I)	Thr (T)	Asn (N)	Ser (S)	C
A	Ile (I)	Thr (T)	Lys (K)	Arg (R)	A
A	Met (M)	Thr (T)	Lys (K)	Arg (R)	G
G	Val (V)	Ala (A)	Asp (D)	Gly (G)	T
G	Val (V)	Ala (A)	Asp (D)	Gly (G)	C
G	Val (V)	Ala (A)	Glu (E)	Gly (G)	A
G	Val (V)	Ala (A)	Glu (E)	Gly (G)	G

Molecular weight of amino acids:

Ala (A)	89 Da
Arg (R)	174 Da
Asn (N)	132 Da
Asp (D)	133 Da
Cys (C)	121 Da
Gln (Q)	146 Da
Glu (E)	147 Da
Gly (G)	75 Da
His (H)	155 Da
Ile (I)	131 Da
Leu (L)	131 Da
Lys (K)	146 Da
Met (M)	149 Da
Phe (F)	165 Da
Pro (P)	115 Da
Ser (S)	105 Da
Thr (T)	119 Da
Trp (W)	204 Da
Tyr (Y)	181 Da
Val (V)	117 Da

	T	C	A	G
Y = pYrimidine	x	x		
W = Weak 3H bonds	x		x	
K = Keto	x			x
M = aMino		x	x	
S = Strong 3H bonds		x		x
R = puRine			x	x
H	x	x	x	
B	x	x		x
D	x		x	x
V		x	x	x
N = aNy base	x	x	x	x



Appendix

Common restriction enzymes

Enzymes	Sequences	Isoschizomers
Apal	GGGCC' C	Bsp120I
Ascl	GG' CGCGCC	PalAI; SgsI
Asel	AT' TAAT	Vspl, AsnI
BamHI	G' GATCC	-
BglII	A' GATCT	-
EagI	C' GGCCG	Eco52I, BstZI, XmaIII, EclXI
EcoRI	G' AATTC	-
EcoRV	GAT' ATC	Eco32I
FspI	TGC' GCA	
HinDIII	A' AGCTT	-
NcoI	C' CATGG	Bsp19I
NdeI	CA' TATG	-
NheI	G' CTAGC	-
NotI	GC' GGCCGC	-
PstI	CTGCA' G	-
Sall	G' TCGAC	
SmaI	CCC' GGG	XmaI, Cfr9I
SnaBI	TAC' GTA	Eco105I
SpeI	A' CTAGT	AcI NI
SphI	GCATG' C	BbuI, PaeI
StuI	AGG' CCT	AatI, Eco147I
XbaI	T' CTAGA	-
XhoI	C' TCGAG	PaeRI



Appendix

Conversions

Spectroscopic extinction coefficients

1 A_{260} unit dsDNA = 50 $\mu\text{g/ml}$

1 A_{260} unit ssDNA = 33 $\mu\text{g/ml}$

1 A_{260} unit ssRNA = 40 $\mu\text{g/ml}$

Centrifugal conversions

$$\text{rpm} = 1,000 * (\text{RCF}/1.12r)^{0.5}$$

$$\text{RCF} = 1.12 * r * (\text{RPM}/1000)^2$$

RCF = relative centrifugal force (x g)

rpm = revolutions per minute

r = radius of rotor in mm

Molecular weight into Dalton

Average molecular weight (MW) of a deoxynucleotide: 330 Da

MW of dsDNA = (number of bp) * 660 Da



Appendix

Gel characteristics

Resolution of linear DNA in agarose gels

Recommended % agarose	Optimum resolution
0.5	1,000 - 30,000 bp
0.7	800 - 12,000 bp
1.0	500 - 10,000 bp
1.2	400 - 7,000 bp
1.5	200 - 5,000 bp
2.0	50 - 2,500 bp

Resolution of proteins in polyacrylamide gels

Recommended % acrylamide	Protein size range
8	40 - 200 kDa
10	21 - 100 kDa
12	10 - 40 kDa



Appendix

Proteins

Fusion proteins and tags

Names	Residues (aa)	Molecular weight (kDa)
FLAG-tag	8	1.0
Glutthion-S-transferase (GST)	211	26
Green fluorescent protein (GFP)	238	26.9
HA-tag	9	1.2
His-tag	6, 8 or 10	0.8, 1.1 or 1.4
Maltose-binding protein (MBP)	369	40
Myc-tag	10	1.2
S-tag	15	1.7
Strep-tag	8	1
T7-tag	11	1.1

Protease cleavage sites

- Enterokinase: (N/D) (N/D) (N/D) (N/D) K↓x
- Factor Xa: (A/F/G/I/L/T/V/M) (E/D) GR↓x;
most suitable: (A/I) (E/D) GR↓x
- TEV: ExxYxQ↓ (G/S) ;
most common: ENLYFQG
- Thrombin: LVPR↓ (not D/E) (not D/E)
- Trypsin: (R/K) ↓ (not P) ;
also WK↓P
also MR↓P
not (C/D) K↓ (D/H/Y) ;
not CR↓K;
not RR↓ (H/R)



Appendix

Entelechon sequencing primers

BGHrev	TAGAAGGCACAGTCGAGGC
CMV fwd	CGCAAATGGGCGGTAGGCGTG
GEX5´for	GGGCTGGCAAGCCACGTTTGGTG
GEX3´rev	GGAGCTGCATGTGTCAGAGG
GL2	CTTTATGTTTTTGGCGTCTTCCA
M13for	GTAAAACGACGGCCAG
M13rev	CAGGAAACAGCTATGAC
pECFP-C1-for	CCTGAGCAAAGACCCCAACG
pECFP-C1-rev	CATTTTATGTTTCAGGTTTCAGG
pEGFP-N1-fwd	GGTTTAGTGAACCGTCAGATCCG
pEGFP-N1-rwd	ACCACCCCGGTGAACAGCTCCTC
pECFP-N1-rev	CTTGCTCACCATGGTGGC
pEYFP-N1-for	GGGAGTTTGTGTTTGGCACC
pEYFP-N1-rev	CTCGCGGGACACGCTGAAC
pLenti-rev	CTTTCCACACCCTAACTGACACACATT
pQC5´-fwd	ACGCCATCCACGCTGTTTTGACCT
pQC3´-rwd	AAGCGGCTTCGGCCAGTAACGTTA
pQEprom	CGGATAACAATTTACACACAG
pQEterm	GTTCTGAGGTCATTACTGG
pREP4_for	CGCAAGGGCTGCTAAAGGAA
pREP4_rev	TTCCCGCTTCAGTGACAACG
pUC18for	CTATTACGCCAGCTGGCGAAAG
pUC18rev	GTTAGCTCACTCATTAGGCAC
pVL1392for	AAAATGATAACCATCTCGC
pVL1392rev	GTCCAAGTTTCCCTGTAG
RV3	CTAGCAAAATAGGCTGTCC
RV4	GACGATAGTCATGCCCCGCG
SP6	ATTTAGGTGACACTATAGAA
T3	ATTAACCCTCACTAAAGGGA
T7prom	TAATACGACTCACTATAGGG
T7term	GCTAGTTATTGCTCAGCGG



Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com





Contact data

Entelechon GmbH
Industriestr. 1
93077 Bad Abbach
Germany

Tel. +49 (9405) 96 999 10
Fax +49 (9405) 96 999 28

contact@entelechon.com

www.entelechon.com

