Optimisation of a transcription-translation coupled *in vitro* system

Inaugural-Dissertation zur Erlangung des Doktorgrades des Fachbereichs Biologie, Chemie und Pharmazie der Freien Universität Berlin

Vorgelegt von
Madina Iskakova
aus Almaty, Kasachstan
2005

Gutachter:
Herr Prof. Dr. Knud H. Nierhaus
Herr Prof. Dr. Mathias Ziegler

Datum der Dissertation:
25. Mai 2005

Summary

Cell-free protein synthesis exploits the catalytic machinery of the cell to produce active proteins. An *in vitro* system is flexible and well controlled, and it offers several advantages over conventional *in vivo* technologies such as easy ways for purification, synthesis of regulatory and/or toxic proteins, incorporation of artificial or modified amino acids that might be doted with isotopes required for NMR.

Here I describe experiments exploring optimisation possibilities concerning yield and quality of the synthesised protein. Some experimental strategies also include expression of eukaryotic genes in prokaryotic expression systems.

The following results have been achieved:

1: Quality criteria developed that allow a critical evaluation of parameters important for the coupled transcription/translation system or improving the yield and quality of the synthesized protein exploiting the features of the green fluorescent protein GFP.

2: The standard transcriptase used in overexpression studies *in vivo* and *in vitro* is the T7 polymerase. The fundamental difficulty with this enzyme is the fact that it is about six times faster than the *E. coli* transcriptase and thus uncouples transcription from translation, a possible reason for the fact that *in vitro* systems usually produce proteins with an activity of 30 to 60% only. We tested some slow mutants of T7 polymerase that approached the rate of the *E. coli* transcriptase and observed indeed a significant improvement up to 100% of the active fraction, although at the cost of lower yields.

3: A similar improvement of the active fraction was observed at lower incubation temperatures down to 20°C, again at the cost of lower yields.

4: According to literature data some amino acids are metabolised during *in vitro* incubations and thus could cause a limitation of protein synthesis. Indeed, we demonstrate that a second addition of amino acids in the middle of the incubation triggers a burst of further protein synthesis. Using this trick at 20°C

pushed the yield of protein to almost that seen at 30°C, but now with an active fraction of 100%. In contrast, our analysis revealed that NTPs are not limiting the gene expression *in vitro* in our system (modified Roche RTS).

5: It is known that the codon usage of highly and lowly expressed proteins in *E. coli* differs dramatically. When we examined this point with human genes, to our surprise a corresponding difference could not be observed. Due to this fact it was possible to identify 11 tRNAs the corresponding codon are quite often used in human genes but rarely in *E. coli* genes. Therefore, for a good expression of eukaryotic genes in *E. coli* systems these 11 tRNAs should be added (and not only the 7 tRNAs supplied in systems from Novagen).

6: I outlined some ways to improve further the expression system.

Zusammenfassung

Zellfreie Proteinsynthese benutzt den Translations- und manchmal auch den Transkriptionsapparat der Zelle zur Synthese aktiver Proteine. Ein *in vitro* System ist flexibel und gut kontrollierbar, und es birgt zahlreiche Vorteile gegenüber *in vivo* Techniken, Beispiele sind Synthese von regulativen und/oder toxischen Proteinen, Einbau von artifiziellen oder modifizierten Aminosäuren, die sogar mit seltenen Isotopen für NMR Untersuchungen dotiert sein können.

In dieser Arbeit beschreibe ich Experimente, die Optimierungs-möglichkeiten bezüglich Ertrag und Qualität der synthetisierten Proteine untersuchen. Einige Experimentalstrategien beziehen sich auch auf die Expression eukaryontischer Gene in prokaryontischen Systemen.

Folgende Ergebnisse wurden erreicht:

1: Qualitätskriterien wurden entwickelt, die eine präzise Bestimmung der synthetisierten Proteinmenge als auch die aktive Fraktion zu messen gestatten. Dabei wurden die Eigenschaften des grün-fluoreszierenden Proteins GFP ausgenutzt.

2: Als Standard-Transkriptase wird *in vivo* als auch *in vitro* T7 Polymerase benutzt. Die grundsätzliche Schwierigkeit mit dieser Transkriptase beruht darauf, dass sie etwa sechs-mal schneller ist als die *E. coli* Transkriptase und damit Transkription von der Translation entkoppelt. Das ist vermutlich der Grund, weshalb *in vitro* Systeme Proteine produzieren, die nur zu 30 bis 60% aktiv sind. Wir testeten einige langsame T7 Polymerasen-Mutanten und beobachteten tatsächlich eine Verbesserung bis zu 100% der aktiven Fraktion, allerdings auf Kosten des Ertrags.

3: Eine ähnliche Verbesserung der aktiven Fraktion wurde bei verminderter Inkubationstemperatur (bis 20°C) beobachtet, jedoch wieder auf Kosten der synthetisierten Menge.

4: Von Literaturdaten wissen wir, dass einige Aminosäuren während der Inkubation metabolisiert werden, was die Proteinsynthese limitieren könnte.

Tatsächlich konnten wir zeigen, dass eine zweite Zugabe von Aminosäuren in der Mitte der Inkubation einen dramatischen Schub der Proteinsynthese auslöste. Die Kombination dieses Tricks mit einer Inkubation bei 20°C vermehrte die Proteinmenge zu der, die bei 30°C gefunden wurde, jetzt aber mit 100% aktiver Fraktion.

5: Es ist bekannt, dass der Codon-Gebrauch bei hoch und niedrig exprimierten Proteinen in *E. coli* deutlich unterschiedlich ist. Eine entsprechende Untersuchung von hoch und niedrig exprimierten Genen im menschlichen Genom offenbarte zu unserer Überraschung keinen unterschiedlichen Codongebrauch. Deshalb war es möglich, 11 tRNA anzugeben, deren zugehörige Codone recht häufig in eukaryontischen mRNAs anzutreffen sind, aber selten in *E. coli* mRNAs vorkommen. Diese 11 tRNAs sollten zur *E. coli* Gesamt-tRNA zugegeben werden, um eine optimale Expression eukaryontischer mRNA in *E. coli* Systemen zu gewährleisten. Unser Ergebnis kontrastiert zu den 7 tRNAs in Zahl und Art, die von Novagen angegeben werden.

6: Ich gebe schließlich zusätzliche Hinweise zu einer weiteren Optimierung der *in vitro* Proteinsynthese.

Table of Contents

Abbreviations

AA	acrylamide
aa-tRNA	aminoacyl-tRNA
Å	Angstrom
Asn	asparagine
Asp	aspartic acid
ATP	adenosine triphosphate
BAA	bis-acrylamide
BPB	bromphenol blue
BSA	bovine serum albumine
cpm	counts per minute
Da	dalton
DNA	desoxyribonucleic acid
dsDNA	double strand DNA
DTT	dithiothreitol
EF-G	elongation factor G
EF-Ts	elongation factor thermo stable
EF-Tu	elongation factor thermo unstable
F.I.	fragmentation index
J	joule
GDP	guanine diphosphate
GFP	green fluorescent protein
Gly	glycine
GTP	guanine triphosphate
His	histidine
HPLC	high performance liquid chromatography
HSS	HEPES salts solution
HSWP	high-salt washed proteins

IF	initiation factor
IPTG	isopropyl-beta-D-thiogalactoside
IRES	internal ribosomal entry site
k-	kilo-, 1000
kb	kilobase, 1,000 nucleotide bases
Leu	leucine
m-	milli-, 0.001
mA	milliampere
µ	micro-, 0.000001
µCi	microcurie
M-	mega-, 1,000,000
MDa	megadaltons
MgAc	magnesium acetate
MQ	milliQ water
mRNA	messenger RNA
M.W.	molecular weight
NaAc	sodium acetate
NH_4Ac	ammonium acetate
nt	nucleotide(s)
NTP	nucleoside triphosphate
Ω	Ohm
PAA(G/gel)	polyacrylamide gel
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEP	phosphoenol pyruvate
PK	pyruvate kinase
Phe	phenylalanine
P.I.	protease inhibitors
Poly(A)	poly-adenine mRNA
Poly(U)	poly-uridine mRNA

PPi	inorganic pyrophosphate
PTF	peptidyl transferase centre
RBS	ribosomal binding site
RF	release factor
RNA	ribonucleic acid
RNAP	RNA polymerase
RNase	ribonuclease
rpm	revolutions per minute
rRNA	ribosomal RNA
RRF	ribosome recycling factor
RTS	rapid translation system
S	Svedberg unit (sedimentation coefficient)
SCS	small compound solution
SD	Shine-Dalgarno sequence
SDS	sodium dodecylsulphate
TCA	trichloroacetic acid
tRNA	transfer RNA
Tyr	tyrosine
UTR	untranslated region
V	volt
v/v	volume/volume
WT	wild type
w/v	weight/volume

Acknowledgements

I am glad to thank Prof. Knud Nierhaus for giving me the opportunity to work in his group, for his supervision and kind concern. I am very thankful for the knowledge I have gained on the ribosome during the timeframe of my thesis.

I would like to thank the entire group, who are more than colleagues to me and also for their kind collaboration I thank Detlev, Edda, Oliver, Pavel, Viter, Witold, Yan, and Yoshika. Together with Witold we were working on the optimisation of the *in vitro* systems and sharing the tricks we learned. I thank him for his participation and comments on the results (sections 3.6.2 and 3.6.3).

I am grateful to all other ribosomologists on our floor, to Daniel, Joerg, Klaus, Paola, Sean and Uschi.

I am glad that Evelyn, Tanya and Pasha, Alix, Diana, Julie and Handan visited our laboratory. My special gratitude for their visits to Andrei, Olga and Peter; who are coming from the laboratory of Prof. Dr. Olga Dontsova in Moscow.

I am thankful to Gabriele for her help and advices from the very beginning of my arrival to Berlin. Thankfulness to librarians for their attention in providing the required literature during the work frames of my thesis. I also would like to thank all people at the workshop.

I am grateful to all my friends and neighbours in the fraternity "Turnerschaft Berlin im CC" for their hospitality and kind understanding.

A very special thank is to my parents who cheered me a lot, and to my younger brother Tair for his joyfulness. I was very happy that they could visit me in Berlin.

CV OF MADINA ISKAKOVA

Max-Planck-Institut für Molekulare Genetik

AG Ribosomen, Prof. Knud H. Nierhaus

Ihnestr. 73 Tel.: + 49 – 8413 – 1588

D-14195 Berlin, Germany Fax: + 49 – 8413 – 1690

iskakova@molgen.mpg.de

PERSONAL INFORMATION

- Birth date: Jan. 14, 1979
- Birthplace: Alma-Ata, USSR
- Citizenship: Republic of Kazakhstan
- Fluent in Russian and in English, proficient in German

EDUCATION

- Ph. D. in Molecular Biology. Under the supervision of Prof. Knud H. Nierhaus. Dissertation entitled "Optimisation of a transcription-translation coupled *in vitro* system." Max-Planck-Institute for Molecular Genetics, Berlin, Germany, May 2005.
- M. D. in Biotechnology. Thesis title: "Cloning and structural and functional analysis of potato virus Y genomic RNA 3'-untranslated region." Kazakh National State University of Al-Farabi, Almaty, Kazakhstan, July 2001.
- Training Course in Biodiversity organised for six months by the United Nations University in Ghent University, Ghent, Belgium, May 2000.
- B. S. in Biotechnology. Kazakh National State University of Al-Farabi, Almaty, Kazakhstan, July 1999.
- Secondary school No. 120 with deep study of English, Almaty, Kazakhstan, June 1995.