Question 7: Optimized Energy Consumption for Protein Synthesis

Witold Szaflarski · Knud H. Nierhaus

Abstract In our previous contribution (Nierhaus, Orig Life Evol Biosph, this volume, 2007) we mentioned that life had solved the problem of energy supply in three major steps, and that these steps also mark major stages during the development of life. We further outlined a possible scenario concerning a minimal translational apparatus focusing on the essential components necessary for protein synthesis. Here we continue that consideration by addressing on one of the main problems of early life, namely avoiding wasteful energy loss. With regard to the limiting energy supply of early living systems, i.e. those of say more than 3,000 Ma, a carefully controlled and product oriented energy consumption was in demand. In recent years we learned how a bacterial cell avoids energy drain, thus being able to pump most of the energy into protein synthesis. These lessons must be followed by the design of a minimal living system, which is surveyed in this short article.

Keywords Minimal living systems · Protein synthesis · Optimized energy consumption · Idle ribosomes · EF-G

Introduction

Protein synthesis is an energy demanding process: Calculating the energy required for (1) the synthesis of a codon, which is the nucleic acid information unit for an amino acid, (2) the charging reaction of a tRNA by its synthetase with the correct (cognate) amino acid, and (3) the subsequent incorporation of this amino acid into the nascent peptide chain, a total of 10 energy-rich bonds need to be sacrificed. In this context an energy rich bond means an acidic-anhydride bond of adjacent phosphate residues of ATP or GTP, each with an energy content of about $\Delta G^{0\prime} = -6$ kcal/mol. The enormous energy requirement explains why a cell has developed intricate systems for controlling energy consumption. This is precisely the

W. Szaflarski · K. H. Nierhaus (✉)
Max-Planck-Institut für Molekulare Genetik, AG Ribosomen, Ihnestrasse 63-73, 14195 Berlin, Germany
e-mail: nierhaus@molgen.mpg.de

Achilles' heel of modern in vitro translation systems, where protein synthesis is severely restricted by a huge squandering of energy: Usually no more than 5% of the energy is used for actual synthesis of proteins, with the rest wasted by uncontrolled and useless energy drains. In striking contrast, a bacterial cell can channel up to 70% of its energy consumption into protein synthesis. In the following sections we will outline the major points one has to follow in order to avoid energy losses and to direct most of the energy – almost like the living cell – into useful work, in this case, protein synthesis.

At least three major reasons for energy losses in vitro systems can be identified: (1) the metabolic energy drain, (2) idle ribosomes and an excess of the elongation factor EF-G, and (3) a metabolization of amino acids. In this order we consider these points.

Preventing Metabolic Energy Drain

The standard way to regenerate energy (ATP, GTP) used in protein synthesis is the addition of millimolar concentrations of phosphoenolpyruvate (PEP) and the enzyme pyruvate kinase that transfers the phosphate group from PEP to AMP and ADP yielding ADP and ATP, respectively (see Fig. 1a). Likewise, GMP and GDP are also substrates for this
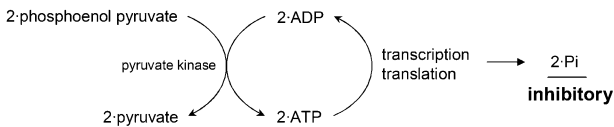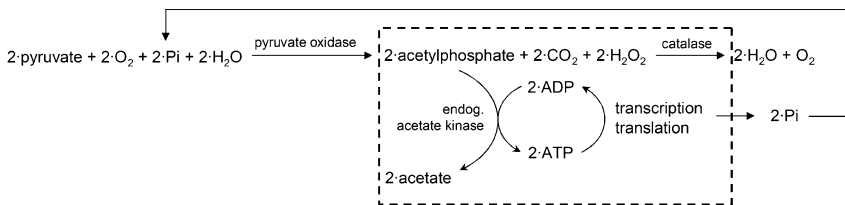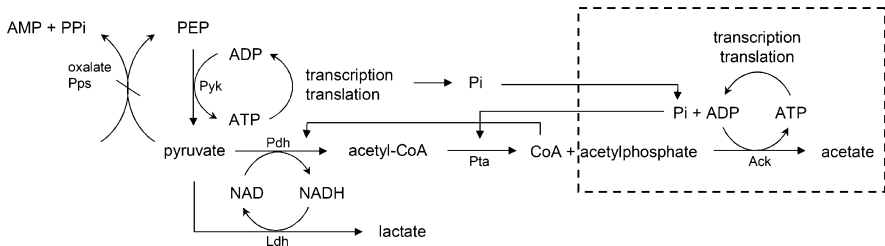

Fig. 1 Various ATP regeneration systems for in vitro protein expression. **a** Conventional scheme for energy regeneration using phosphoenol puryvate (*PEP*) and pyruvate kinase. The generated orthophosphate (*Pi*) inhibits protein synthesis. **b** Improved energy regeneration using pyruvate, pyruvate oxidase and endogenous acetate kinase (according to Kim and Swartz 2000). **c** Near in vivo system based on phospho*trans*acetylase (*Pta*) to produce acetylphosphate. The *boxes* in **b** and **c** show a common set of reactions for regeneration of *ATP*. The PEP synthase (*Pps*) is blocked by the presence of oxalate. *Ack* acetate kinase, *Ldh* lactate dehydrogenase, *Pdh* pyruvate dehydrogenase, *Pyk* pyruvate kinase. According to Jewett and Swartz (2004), modified

enzyme. PEP contains the energy richest bond of all biological samples, the phosphoester bond has an energy content of $\Delta G^{0\prime}=-12$ kcal/mol. This system allows only the synthesis of limited amounts of proteins due to the inhibitory effect resulting from the generation of orthophosphate. Orthophosphate byproducts, not only reduce the pH, but also bind Mg^{2+} ions, thus reducing the free Mg^{2+} concentration and severely effecting the structure of all ribonucleoproteins in the system, the most prominent one being the ribosome. The two positive valences of Mg^{2+} shield the negatively charged phosphate groups of the RNA and maintaining the high density of the ribosome, thus reducing the free concentration of Mg^{2+} sets the repelling phosphate groups free and expands the ribosome, impairing its efficiency.

One way to get rid of this unwanted effect is to recycle the liberated phosphate by addition of pyruvate oxidase, together with some prosthetic groups, such as TPP and FAD (thiamin pyrophosphate and flavin dinucleotide, respectively). This system generates acetylphosphate, a standard component of the regeneration system of a cell, by using molecular oxygen and H_2O (Fig. 1b), and this improvement in conditions leads to a definite prolongation in protein synthesis in vivo (Kim and Swartz 2000).

This system, however, is still far from an ideal situation. The need of molecular oxygen is the drawback, since the uptake of oxygen is limited and becomes worse when larger reaction volumes are used because of the decreasing relative surface area. Jewett and Swartz (2004) recently reported a convincing solution of this problem: by combining the classical PEP/pyruvate kinase system with the synthesis of acetylphosphate via acetyl-CoA, they efficiently recycle orthophosphate (Pi) during the synthesis of acetylphosphate (Fig. 1c). This solution represents a breakthrough: for the first time several mg of proteins can be synthesized per one ml, making in vitro systems a serious alternative for synthesis of proteins that can be used for structural and functional analyses. Table 1 compiles the precise

Table 1 Final concentrations of components in an in vitro system for protein synthesis

Component	Final concentrations in our reaction mix
HEPES–KOH (pH 8.2)	60 mM
Ammonium acetate	80 mM
Potassium glutamate	230 mM
Sodium oxalate	3 mM
DTT	2 mM
Cycle-AMP	0.7 mM
Folinic acid	35 μg/ml
tRNAs	350 μg/ml
NADH	0.35 mM
Coenzyme A	0.3 mM
ATP	1.5 mM
CTP	1 mM
GTP	1 mM
UTP	1 mM
PEG-8000	2% (w/v)
20 amino acids	2 mM each
PEP	35 mM
Magnesium acetate	12 mM
T7 RNA polymerase	100 μg/ml
E. coli S30 cell lysate	4–6 A_{260}
Plasmid	4 μg/60 μl
Rifampicin	10 μg/ml

conditions that allow the synthesis of mg of a protein in a coupled transcription–translation system, where the gene of interest is introduced on a plasmid behind a T7 promoter (for details see Iskakova et al. 2006).

Prevent Idle Ribosomes in the Presence of an Excess of EF-G

GTP-binding (G) proteins are an important class of regulatory proteins. They bind GDP or GTP, and undergo a conformational change depending on the presence of which nucleotide is bound. In the GTP conformer (the "on"-state) they bind to their target and promote a distinct reaction of the target, with the energy being paid by the binding energy generated during the binding step of the G protein. The target then signals that the reaction was successfully accomplished and triggers the hydrolysis of "GTP→GDP+Pi" in the corresponding enzymatic center of the G protein. Upon release of the inorganic phosphate, the G protein switches into the GDP conformer ("off"-state) and looses it's affinity for the target and thus dissociates from it. Four G proteins participate in translation, two of which are components of the central activity of the ribosome, the elongation cycle, where in a cycle of reactions the nascent chain is prolonged by one amino acid (aa). One factor is EF-Tu, which brings the aa in the form of aa-tRNA to the decoding center of the ribosome, the other is EF-G, which translocates the tRNAs on the ribosome by one codon length (reviewed by Wilson and Nierhaus 2003).

Usually a G protein recognizes a specific configuration in the target before promoting or even triggering a specific reaction. This is also true for EF-G; it recognizes specifically the ribosome state before translocation and dissociates from the ribosome following GTP cleavage and Pi release. However, EF-G is the only G protein involved in translation that also interacts effectively with empty or idle ribosomes, probably because empty ribosomes seem to have "conformational memory" of the two main states before and after translocation (Mesters et al. 1994). Thus empty ribosomes can trigger EF-G dependent GTP hydrolysis with a high turnover (Fig. 2), which is not seen with EF-Tu. It follows that in a minimal system with poor energy resources both ribosomes and EF-G must be present in limiting amounts. The latter is established, if the molar ratio of EF-G: ribosomes is about 0.2 to 0.3, i.e. 2–3 EF-G molecules per 10 ribosomes.

Considerations on the Metabolization of Particular Amino Acids (Arg, Cys, Ser and Trp)

When we follow synthesis of the Green Fluorescent Protein (GFP) kinetically in an efficient coupled transcription–translation system, synthesis occurs almost linearly for about seven

Fig. 2 Kinetics of EF-G dependent hydrolysis of GTP on 70S ribosome. Taken from Qin et al. 2006, where the experimental details can also be found
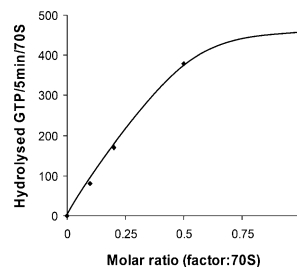
Fig. 3 Synthesis of GFP in a near in vitro protein system. **a** A comparison of the synthesis of GFP and that of GFP-mRNA. *Continuous line*, total synthesis of GFP determined by SDS PAGE; *dashed line*, amount of GFP-mRNA as determined by Northern blotting. **b** Total synthesis of GFP; (−)aa, 2 mM amino acids present from the beginning; (+)aa, a second addition of amino acids after 7 h incubation leading to a total amino acid concentration of 4 mM. Taken from Iskakova et al. (2006), see also for experimental details
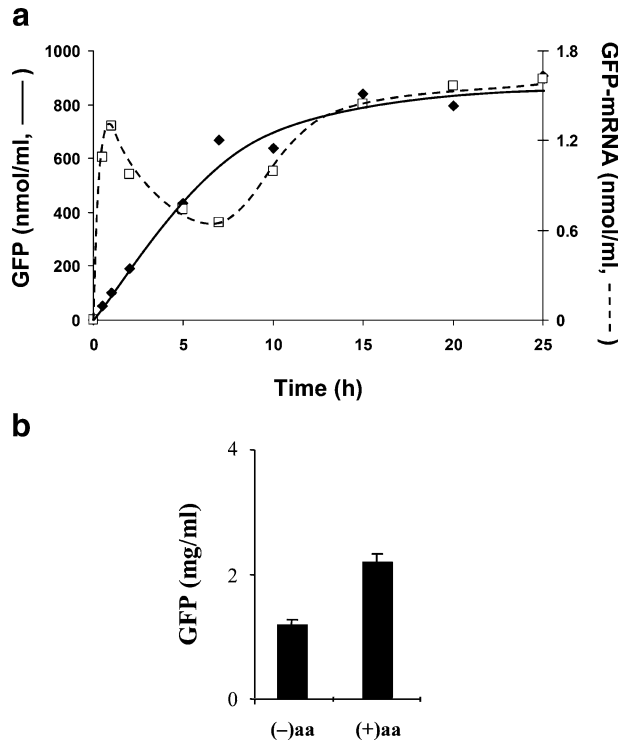
hours, but then starts to plateau (Fig. 3a). One reason for this plateau could be consumption of the energy-rich compounds (e.g. ATP). However, when we simultaneously follow the synthesis of mRNA in the same reaction, we observe that the total mRNA (synthesis and decay) displays a minimum at 7 h, but as GFP synthesis ceases, the concentration of mRNA starts to increase again. This clearly indicates that there is still enough energy in the system. Another possibility therefore could be the metabolic degradation of some amino acids, which would block protein synthesis completely. Indeed, some amino acids have been identified that are significantly metabolized in vitro systems, in particular arginine, cysteine, serine and tryptophan (Michel-Reydellet et al. 2004); for review see (Swartz 2006). Consistent with this idea, refreshing the amino acid pool after a 7-h incubation period significantly improved the overall yield (Fig. 3b).

Conclusions

Together with our article in chapter 6.3, we have tried to estimate the minimal set of genes needed for a minimal cell. A brave estimate of the minimal components for the translational apparatus today comprises no more than 200 genes, of which more than 120 are associated with the translational apparatus, encoding about 40 genes for ribosomal proteins, two rRNAs (omitting the 5S rRNA), 21 tRNAs, 20 synthetases, six factors and at least 20 tRNA modifying enzymes. In addition, a minimum of 30 genes are needed for both the household energy as outlined in this article and the synthesis of at least some of the amino acids (note that some of the amino acids were formed in the Stanley Miller experiments mimicking the

atmosphere and the physical environment more than 3 Bya (Miller 1953) and therefore need not be synthesized by the cell).

Even if we consider a "limping" life form, we are ending up with a total of at least 150 genes. This is in a reasonable agreement with the most restricted life form known today, namely the bacteria *Carsonella rudii* with 189 genes (Nakabachi et al. 2006). This bacteria is an obligatory endosymbiont living in the "jumping plant lice" (psyllids), namely in specialized cells (bacteriocytes) made by the insects. The bacteria belong to the γ-proteobacteria as *Escherichia coli*, are transmitted by vertical transmission through the host generations (no exogenous infections!) and contain a full translation system as well as prepare some amino acids for the host. We think therefore that it is not possible at the moment to construct a primitive cell "getting-by" with a number of genes significantly below the 150 genes mentioned above.

References

Iskakova MB, Szaflarski W, Dreyfus M, Remme J, Nierhaus KH (2006) Troubleshooting coupled in vitro transcription–translation system derived from *Escherichia coli* cells: synthesis of high-yield fully active proteins. Nucleic Acids Res 34:e135

Jewett MC, Swartz JR (2004) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. Biotechnol Bioeng 86:19–26

Kim DM, Swartz JR (2000) Prolonging cell-free protein synthesis by selective reagent additions. Biotechnol Prog 16:385–390

Mesters JR, Potapov AP, de Graaf JM, Kraal B (1994) Synergism between the GTPase activities of EF-Tu. GTP and EF-G.GTP on empty ribosomes. Elongation factors as stimulators of the ribosomal oscillation between two conformations. J Mol Biol 242:644–654

Michel-Reydellet N, Calhoun K, Swartz J (2004) Amino acid stabilization for cell-free protein synthesis by modification of the *Escherichia coli* genome. Metab Eng 6:197–203

Miller SL (1953) A production of amino acids under possible primitive earth conditions. Science 117:528–529

Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, Hattori M (2006) The 160-kilobase genome of the bacterial endosymbiont *C*arsonella. Science 314:267

Nierhaus KH (2007) Early steps of evolution and some ideas about a simplified translational machinery. Orig Life Evol Biosph (this volume)

Qin Y, Polacek N, Vesper O, Staub E, Einfeldt E, Wilson DN, Nierhaus KH (2006) The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. Cell 127:721–733

Swartz J (2006) Developing cell-free biology for industrial applications. J Ind Microbiol Biotechnol 33:476–485

Wilson DN, Nierhaus KH (2003) The ribosome through the looking glass. Angew Chem Int Ed Engl 42:3464–3486