

Proto-Organism Kinetics: Evolutionary Dynamics of Lipid Aggregates with Genes and Metabolism

Steen Rasmussen^{†,‡}, Liaohai Chen[‡], Bärbel M. R. Stadler^{§,¶}
and Peter F. Stadler^{#,§,‡}

[†]*EES-6 MS T003 & T-CNLS MS B258, Los Alamos Natl. Lab., Los Alamos NM 87545, USA*

[‡]*Santa Fe Institute, Santa Fe, NM 87501, USA*

[‡]*Bioscience Division, Argonne Natl. Lab., Argonne, IL 60439, USA*

[§]*Inst. f. Theor. Chemistry & Structural Biol., Univ. Vienna, Austria*

[¶]*Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany*

[#]*Bioinformatik, Inst. f. Informatik, Univ. Leipzig, Germany*

(steen@lanl.gov, lhchen@anl.gov, stadler@mis.mpg.de, studla@tbi.univie.ac.at)

2002/09/26

Abstract. A synthetic proto-organism could be self-assembled by integrating a lipid proto-container with a proto-metabolic subsystem and a proto-genetic subsystem. This three-component system can use energy and nutrients either by means of redox or photo-chemical reactions, evolve its proto-genome by means of template directed replication, and ultimately die. The evolutionary dynamics of the proto-organism depends crucially on the chemical kinetics of its sub-systems and on their interplay. In this work the template replication kinetics is investigated and it is found that the product inhibition inherent in the ligation-like replication process allows for coexistence of unrelated self-replicating proto-genes in the lipid surface layer. The combined catalytic effects from the proto-genes on the metabolic production rates determine the fate of the strain proto-cell.

1. Introduction

In the last years at least two proposals for artificial minimal cells have been put forward [10, 17]. Both call for a rather sophisticated molecular machinery to be enclosed in a lipid vesicle. The model of Szostak *et al.* consists of a vesicle containing an RNA genome that contains an RNA-replicase ribozyme (e.g. an advanced version of the molecule described in [6]) and a functionality that influences the fitness of the vesicle. The construct of Pohorille & Deamer, which is even closer to a modern cell, includes transcription and translation functionalities.

In this contribution we investigate the consequences of embedding very simple genetic material in lipid aggregates that actively facilitates an auto-catalytic reproduction of lipids as well as the genetic material itself. Our design is inspired by original concepts for a protocell by Luisi *et al.* [7], but our model also differs in several respects from this and other proposals: (1) Our focal point is a minimalistic, thermodynamic coupling between the three functional structures container, metabolism, and genes. (2) Instead of RNA we envision simpler molecules, such as peptide nucleic acids (PNA)

[9], that may be much easier to couple with the lipid layer than RNA due to its hydrophobic backbone. (3) Instead of sophisticated ribozymes we consider very simple short oligos that are capable of enzyme-free self-replication by means of a ligation mechanism [20]. (4) As in the other protocell proposals we utilize the lipid to keep the cooperative structure together although the proto-genetic activity is not on the inside of a vesicle, e.g. [1], but on the outside of a lipid aggregate. We can therefore work with simpler lipid structures such as micelles [23]. (5) We make explicit use of the differences of the thermodynamic properties within the lipid phase compared to the water and the lipid/water interface. As a result we obtain a quite different chemistry.

2. The Los Alamos Bug

The Los Alamos Bug is a proposed synthetic “organism” that integrates three functionalities and their associated molecular aggregates: (i) a lipid container, (ii) a photo-metabolic system, and (iii) a hydrophobically anchored templating polymer that influences metabolic kinetics, Figure 1. We focus on the main conceptual design here, as the physicochemical and thermodynamic issues involved in the assembly of this proto-organism are discussed in detail in [11]. The main assembly steps are the integration of a simple metabolic system in an amphiphilic aggregate, either a micelle or a vesicle bilayer, (l.h.s. of Fig. 1) and the integration of a templating polymer in an amphiphilic phase (r.h.s. of Fig. 1). The combination of these two subsystems defines a very simple proto-organism.

Amphiphiles (lipids) self-assemble into micelles or vesicles that can host templating polymers with a hydrophobic backbone which remain anchored in the lipid aggregate because it is thermodynamically favorable. For two reasons we consider PNA or PNA-like nucleic acid analogues as the primitive genetic material: The hydrophobic backbone makes a strong interaction with the lipid phase plausible and PNA is more plausible than RNA or DNA in terms of prebiotic synthesis [8]. For simplicity we will use PNA as a representative of such a class of biopolymers without insisting on a particular backbone chemistry or a particular nucleic acid alphabet.

In one possible physico-chemical implementation we assume visible light as the primary energy source. An organic photo-sensitizer is added to enhance the efficiency of photo electron transfer which is hydrophobic and hence sinks into the lipid aggregate. Non-functional precursors of the lipids and the templating polymer have to be available as “food molecules”. The precursors themselves cannot form micelles or vesicles or act as template, resp., they are hydrophilic, however, so that they will readily associate with the lipid aggregate. For a discussion of photochemical systems see [19, 4].

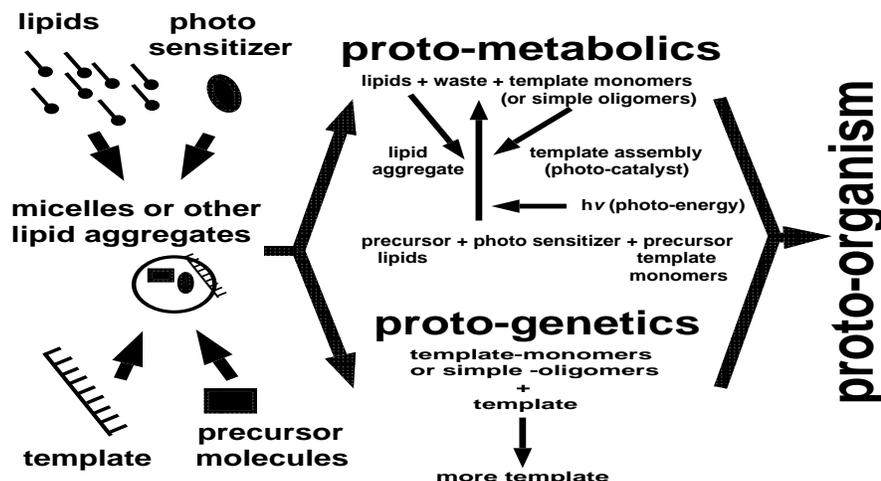


Figure 1. Two experimental systems are combined to form a simple proto-organism as a templating polymer with a hydrophobic backbone is associated with a lipid aggregate. Left: self-assembly of the different components. Middle top: photo active reactions (photo-fragmentation) within the lipid aggregate using a particular nucleobase sequence as an electron relay system that enhances the quantum yield of the photo reactions. Middle bottom: a set of templating directed ligation and replication reactions at the water/lipid interface as well as within the interior of the lipid aggregate. See text and [11] for details.

These precursor molecules are chemically modified to functional lipids and templates. For example, both the lipid precursor and the PNA precursors can have hydrophobic groups attached via an ester bond that has to be broken by the metabolic reaction. The energy for this reaction is taken from the charge separation triggered by the capture of light energy through the photo-sensitizer. The quantum yield for breaking the ester bond is not very high, however. Adding a charge transfer chain to the system can help to channel the positive charge away from the sensitizer molecule very quickly, thereby arresting the excitation of the precursor and forcing the bond fragmentation. A key property of this particular proto-organism system is that the templating polymer also acts as electron transfer chain and that its electron transfer abilities strongly depend on the specific nucleobase sequence: Only adenine- and guanine-rich sequences have good electron transfer capabilities. This establishes a direct link between metabolic efficiency and genetic information.

3. Replication Kinetics and the Lipid Aggregate Surface

In this section we consider the kinetics of replicating the PNA-like genetic material that is embedded in the membrane of a vesicle or micelle (the reac-

tions on the r.h.s. of Fig. 1. The possibility of polycondensation of aminoacids in the membrane matrix to peptides (a reaction that cannot be obtained in the absence of liposomes) has been well demonstrated [2]. It seems reasonable, hence, to assume that a similar reaction is possible with PNA. A model for the ligation-based replication process is sketched in figure 2: We assume that the single-stranded template is located at the surface, exposing the hydrophilic bases to the aqueous medium while the hydrophobic backbone is sunk into the lipid layer. The system is fed with oligos from the aqueous phase. Partial duplices are therefore obtained by specific hydrogen bond formation. The entire surface of a completed complex ABC is hydrophobic; thus it will sink into the lipid layer. There the ligation step, i.e., the expulsion of H_2O , is thermodynamically significantly more favorable than in water and we eventually obtain the double stranded ligation product CC' . A small fraction of CC' will diffuse back to the surface where it may dissociate into its two strands. The balance between the hybridized CC' complex and the free C and C' strands depends on the nucleo-base binding energies, which in turn depend on the details of the PNA back bone and multiple physicochemical factors (e.g. temperature and salt concentration). The important ingredient in our reasoning is that both C and C' remain anchored at the surface of the lipid layer. This way the “genetic material” is permanently associated with the lipid micelle or vesicle.

Under the assumption of a realistic nucleic acid chemistry we will observe a ligation/replication process that proceeds via complementary copies. In order to simplify the mathematics we consider here a hypothetical direct replication mechanism. Equivalently, we assume that the template is palindromic as in von Kiedrowski’s experiments [13]. Plus-minus complementarity in replication can also be dealt with by assuming stationarity in relative concentrations of plus and minus strand [5]; the plus-minus ensemble can then be treated as a single species [15].

The replication process starts with the formation of partial duplices $A + C \rightleftharpoons AC$ (and analogous for BC) at the surface. In the next step an un-ligated duplex ABC is formed reversibly



The species ABC, however, is hydrophobic and thus quickly sinks into the lipid layer, where the ligation reaction $ABC \rightarrow C_2^*$ takes place. The asterisk indicates that C_2^* is buried in the lipid layer instead of being anchored on the surface. We assume that the ligation step is irreversible. If the ligation step is slow compared to the other reactions (rate limiting step) a selection pressure on a faster ligation reaction will occur.¹ If the ligation step is fast in comparison with the other reactions, then $[ABC] \approx 0$ and the rate of formation of C_2^*

¹ If the ligation step is rate limiting we obtain the same over-all kinetics, equ.(11), albeit with a different dependence of the constants α and β on the elementary rate constants, see [14].

and the surface is fast, imply

$$[AC] = \frac{k_A[A]}{\bar{k}_A + a'[B]}[C] \quad [BC] = \frac{k_B[B]}{\bar{k}_B + a''[A]}[C] \quad [C_2]^* = \frac{\bar{f}}{f}[C_2] \quad (9)$$

$$[C_2] = \frac{\bar{k}_d}{k_d}[C]^2 + \frac{1}{k_d} \left(\frac{a'k_A}{\bar{k}_A + a'[B]} + \frac{a''k_B}{\bar{k}_B + a''[A]} \right) [A][B] \times [C] \quad (10)$$

Equ.(8) can therefore be rewritten in the form

$$\frac{dc}{dt} = \alpha c \psi(\beta c) \quad \text{with} \quad \psi(u) = \frac{2}{u} \left(\sqrt{1+u} - 1 \right) \quad (11)$$

where $\alpha > 0$ is the replication rate in dilute solution and $\beta > 0$ is a measure for the product inhibition. The rate function ψ is monotonically decreasing, satisfies $\psi(u) \rightarrow 1$ for small u and asymptotically behaves like $\psi(u) \sim 2/\sqrt{u}$ for large u . It is straightforward to compute the parameters α and β explicitly in terms of the elementary rate constants. The resulting expression, however, is very complicated and will not be needed in our discussion. In addition to the growth law (11) there will be a degradation reaction or at least a dilution flux $-c\Phi$ that regulates the concentration. In practice this could be implemented by means of serial transfer or a continuously stirred tank reactor. The important observation is that the present model yields the same form of the effective reaction kinetics as Kiedrowski's minimal replicator model [21]. The competition of different templates C_k is therefore described by a replicator equation of the form [24]

$$\dot{x}_k = x_k \left[\alpha_k \psi(c_0 \beta_k x_k) - \sum_j x_j \alpha_j \psi(c_0 \beta_j x_j) \right] \quad (12)$$

where $c_0 = \sum_j c_j$ is the total concentration of replicating material, $x_k = c_k/c_0$ are the relative replicator concentrations, and α_k and β_k are effective kinetic constants that can be computed from the elementary rate constants.

The dynamics of equ.(12) is analyzed in detail in [14]. Most importantly, there is a *survival threshold* $\hat{\Phi}$ such that C_k survives indefinitely if and only if $\alpha_k > \hat{\Phi}$. The threshold fitness $\hat{\Phi}$ depends only on the kinetic constants α_k , β_k , and the total concentration c_0 of replicating material. The parameter $c_0 \beta_k$ describes the importance of product inhibition, which increases with the total concentration c_0 . The global behavior can be tuned by changing c_0 from small to large concentrations. For small c_0 we observe *survival of the fittest*, i.e., only the ‘‘master sequence’’ with the largest value of fitness α_k survives. In this case equ.(12) approaches Eigen's quasispecies model without mutation [5]). For very large concentrations, on the other hand, we have effectively parabolically growing replicators and hence permanent coexistence [16]. In the intermediate regime selection is strong enough to eliminate species with

small values $\alpha_k < \hat{\Phi}$ but not strong enough to reduce the population to the master sequence alone.

The ligation/replication mechanism described here depends upon the propensity of the hydrophobic duplex to bury itself in the lipid phase. Indeed, the expulsion of H_2O from the lipid phase is the thermodynamic driving force in our model. Thus we have to assume that β_k is relatively large and that we effectively operate in a regime where many PNA variants can coexist. We expect product inhibition to become prohibitive for longer chains, hence the size of the individual PNA templates is restricted to very short chains, say, $n \approx 10$ bases.

4. Lipid Aggregate Selection

As we have argued in the previous section, it is possible to accumulate a significant number of different — possibly genetically unrelated — PNA sequences, because the PNA replication on the lipid aggregate surfaces follows an essentially parabolic growth law. This mixture is heritable under the reproduction (i.e., division) of the aggregates since we may assume that the PNAs diffuse freely in and at the surface of the lipid. The vesicles are therefore “tagged” by the PNA genes on their surface which can only propagate within their vesicle and to the off-spring vesicles after division. Even though this would only be true as an approximation, it is sufficient to assume that PNA interchange between aggregates is rare and hence can be treated as a perturbation.

Organizational closure requires an additional coupling between the PNA sequences, or more precisely, between certain PNA sequences and the metabolic production of more lipid molecules and PNA oligomers² from appropriate precursors. This feedback is provided by the sequence dependence of the electron relay efficiency of the PNA which is part of both the genetic templating system and the metabolic reaction complex.

Assuming that the lipid aggregates are either micelles or vesicles we can express the growth of the lipid aggregate in terms of the total water/aggregate area Ω_u ³ of all aggregates with a PNA (mixture) of type u . The autocatalytic formation of vesicle and micelles from their lipid building blocks has been demonstrated with a variety of different amphiphilic molecules [22, 18]. It is reasonable therefore to assume that both the production and the incorporation of new building material is proportional to the surface area, hence

$$\frac{d}{dt}\Omega_u = k_u\Omega_u \quad (13)$$

² PNA is present predominantly in dimer form to avoid cyclization reactions.

³ The same argument holds if Ω defines the hydrophobic volume of the lipid aggregate.

where the growth rate k_u depends explicitly on the catalytic activity of the specific PNA mixture. In particular, aggregates containing only PNA strands poor in G and A will grow slowly or not at all. Off-spring vesicles are viable only if they inherit the photo-catalytically active PNAs.

It is clear that the photo-chemical processes are many orders of magnitude faster than the template directed PNA replication given an excess of precursor molecules. However, an appropriate balance must exist between the production of lipid molecules and the production of PNA oligomers which take part in the template directed ligation (and replication). Many more lipids must be produced for every PNA oligomer to ensure that a larger (or new) lipid aggregate is grown to enable hydrophobic support of a new PNA oligomer. Assuming an appropriate balance between the lipid and PNA oligomer production (say 50 : 1 for micelles and ~ 5000 : 1 for vesicles) the growth law (13) implies that we will observe selection of the fittest PNA/aggregate type.

Since there is an upper and a lower limit on the rate of aggregate reproduction relative to the rate of PNA replication, there is selection for a coupling of PNA replication and aggregate division. If the lipid production is much higher than a balanced production, non-viable aggregates without PNAs are being produced which will be a loss for the system. If the lipid production is too low compared to a balanced rate, the lipid aggregates will be over-loaded with PNAs; this should alter their mechanical properties and presumably will severely disturb replication at the aggregate level, if the chemical system does not break entirely.

Replication of the genes on the surface of the aggregates *per se* does not overcome the limits on the amount of heritable genetic information that is posed by Eigen's error threshold [5]. The growth law (12), however, implies the stability of a genome consisting of a number of tiny "chromosomes" (C_{ks}) if $\alpha_k > \hat{\Phi}$. This is ruled out by the original quasispecies model since it does not allow for coexistence beyond the mutant cloud of the master sequence in homogeneous solution. The possibility of fragmented genomes opens up a much richer scenario for further evolution. Selection can occur at two levels: there is (weak) competition among the "modules" of the genome selecting for template sequences with $\alpha_k > \hat{\Phi}$ and there is (strong) selection for aggregates containing catalytically active PNA strings.

5. Discussion

The protocell discussed here requires much less sophisticated biopolymer components, compared to other proposals [7, 17, 10], and a less well-defined organization of the lipid aggregate. The Los Alamos Bug can therefore be seen as a model for a much earlier step in the transition between nonliving and living matter. If the Los Alamos Bug can indeed be realized experimentally,

it will be endowed with all the prerequisites for combinatorially unlimited Darwinian evolution. Its metabolism will allow it to survive in a “prebiotic” environment although unfavorable environmental conditions (e.g. heat or pH) will kill it. In short, the proto-organism will be alive at least w.r.t. the most common definitions of the term.

The lipid-aggregate proto-organism introduced here might be a starting point to reconcile the Lipid-World model of Lancet & Segré [12] with the necessity of information carrying molecules: dynamically our proto-cells could behave like a GARD, in particular since sequence-dependent catalytic activities of PNA are likely to be comparable to RNA and DNA catalysis. However, the information is not carried by the composition of the lipid membrane but by the catalytically active PNA replicators that are anchored in it.

Finally it should be reiterated that the presented minimalistic thermodynamic coupling between the functional structures of the proto-organism (container, metabolism, and genes) also could be realized using a redox based metabolism that harnesses chemical energy. Perhaps it would even be possible to use RNA instead of PNA as long as amphiphiles with an opposite charge of the RNA backbone are provided.

ACKNOWLEDGEMENTS. Many people have helped us over the years to develop the proto-organism concept including: Shigeaki Abe, Charles Apel, Mark Bedau, Sterling Colgate, Shelley Copley, David Deamer, Gottfried Köhler, Klaus Lackner, Luigi Luisi, John McCaskill, Bernd Mayer, Peter Nielsen, Martin Nilsson, Norman Packard, Andrew Pohorille, Peter Schuster, Lau Sennels, Kolbjorn Tunestrom, David Whitten, Peter Wills, and Daniel Yamins. This work is supported in part by an LDRD grant from the *Center for Space Science and Exploration (CSSE)* at Los Alamos as well as by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Project No. P-13887-MOB.

References

1. Apel, C. L., D. W. Deamer, and M. N. Mautner: 2002, ‘Self-assembled vesicles of monocarboxylic acids and alcohols: conditions for stability and for the encapsulation of biopolymers’. *Biochim. Biophys. Acta* **1559**, 1–9.
2. Blocher, M., D. Liu, P. Walde, and P. L. Luisi: 1999, ‘Liposome-Assisted Selective Polycondensation of α -Amino Acids and Peptides’. *Macromolecules* **32**, 7332–7334.
3. Borghans, J. A. M., R. J. de Boer, and L. A. Segel: 1996, ‘Extending the Quasi-Steady State Approximation by Changing Variables’. *Bull. Math. Biol.* **58**, 43–63.
4. Chen, L., L. Lucia, and D. G. Whitten: 1998, ‘Cooperative Electron Transfer Fragmentation Reactions. Amplification of a Photoreaction through a Tandem Chain Fragmentation of Acceptor and Donor Pinacols’. *J. Amer. Chem. Soc.* **120**, 439–440.
5. Eigen, M.: 1971, ‘Selforganization of Matter and the Evolution of Macromolecules’. *Naturwiss.* **58**, 465–523.

6. Johnston, W. K., P. J. Unrau, M. J. Lawrence, M. E. Glasner, and D. P. Bartel: 2001, 'RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension'. *Science* **292**, 1319–1325.
7. Luisi, P. L., P. Walde, and T. Oberholzer: 1994, 'Enzymatic RNA Synthesis in Self-Reproducing Vesicles: An Approach to the Construction of a Minimal Synthetic Cell'. *Ber. Bunsenges. Phys. Chem.* **98**, 1160–1165.
8. Nelson, K. E., M. Levy, and S. L. Miller: 2000, 'Peptide nucleic acids rather than RNA may have been the first genetic molecule'. *Proc. Natl. Acad. Sci. USA.* **97**, 3868–3871.
9. Nielsen, P. E.: 1993, 'Peptide nucleic acid (PNA): a model structure for the primordial genetic material?'. *Orig. Life Evol. Biosph.* **23**, 323–327.
10. Pohorille, A. and D. Deamer: 2002, 'Artificial cells: prospects for biotechnology'. *Trends Biotech.* **20**, 123–128.
11. Rasmussen, S., L. Chen, M. Nilsson, and S. Abe: 2002, 'Bridging nonliving and living matter'. *Artificial Life*. In press.
12. Segré, D., D. Ben-Eli, and D. Lancet: 2000, 'Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies'. *Proc. Natl. Acad. Sci. USA* **97**, 4112–4117.
13. Sievers, D. and G. von Kiedrowski: 1994, 'Self-replication of complementary nucleotide-based oligomers'. *Nature* **369**, 221–224.
14. Stadler, B. M. R., P. F. Stadler, and P. R. Wills: 2001, 'Evolution in Systems of Ligation-Based Replicators'. *Z. Phys. Chem.* **21-33**, 216.
15. Stadler, P. F.: 1991, 'Complementary Replication'. *Math.Biosc.* **107**, 83–109.
16. Szathmáry, E. and I. Gladkih: 1989, 'Sub-Exponential Growth and Coexistence of Non-Enzymatically Replicating Templates'. *J. Theor. Biol.* **138**, 55–58.
17. Szostak, J., D. Bartel, and P. L. Luisi: 2001, 'Synthesizing life'. *Nature* **409 Suppl**, 387–390.
18. Veronese, A. and P. L. Luisi: 1998, 'An Autocatalytic Reaction Leading to Spontaneously Assembled Phosphatidyl Nucleoside Giant Vesicles'. *J. Amer. Chem. Soc.* **120**, 2662–2663.
19. Volkov, A. G., M. I. Gugeshashvili, and D. W. Deamer: 1995, 'Energy conversion at liquid/liquid interfaces: artificial photosynthetic systems'. *Electrochimica Acta* **40**, 2849–2868.
20. von Kiedrowski, G.: 1986, 'A Self-Replicating Hexadeoxynucleotide'. *Angew. Chem. Int. Ed. Engl.* **25**, 932–935.
21. von Kiedrowski, G.: 1993, 'Minimal Replicator Theory I: Parabolic Versus Exponential Growth'. In: *Bioorganic Chemistry Frontiers, Volume 3*. Berlin, Heidelberg, pp. 115–146.
22. Walde, P., R. Wick, M. Fresta, A. Mangone, and P. L. Luisi: 1994, 'Autopoietic Self-Reproduction of Fatty Acid Vesicles'. *J. Amer. Chem. Soc.* **116**, 11649–11654.
23. Whitten, D. G., L. Chen, H. C. Geiger, J. Perlstein, and X. Song: 1998, 'Self-Assembly of Aromatic-Functionalized Amphiphiles: The Role and Consequences of Aromatic-Aromatic Noncovalent Interactions in Building Supramolecular Aggregates and Novel Assemblies'. *J. Phys. Chem. B* **102**, 10098–10111.
24. Wills, P. R., S. A. Kauffman, B. M. Stadler, and P. F. Stadler: 1998, 'Selection Dynamics in Autocatalytic Systems: Templates Replicating Through Binary Ligation'. *Bull. Math. Biol.* **60**, 1073–1098.

Mailing address for review purposes:

P.F. Stadler, Bioinformatik, Inst. f. Informatik, Univ. Leipzig, Kreuzstraße 7b, D-04103 Leipzig, Germany, Fax: ++49 341 149 5120.

Email: stud1a@tbi.univie.ac.at