From chemical reactions to evolution: emergence of species

Carletti, Timoteo; Fanelli, Duccio

Published in:
Europhysics Letters

Publication date:
2007

Document Version
Early version, also known as pre-print

A Letters Journal Exploring the Frontiers of Physics

January 2007

EPL, **77** (2007) 18005 www.epljournal.org
doi: 10.1209/0295-5075/77/18005

From chemical reactions to evolution: Emergence of species

T. Carletti[1] and D. Fanelli[2,3,4]

[1] *Département de Mathématique, Université Notre Dame de la Paix - 8 Rempart de la Vierge, B-5000 Namur, Belgium*
[2] *Dipartimento di Energetica "S. Stecco", Università di Firenze - via S. Marta 3, I-50139 Firenze, Italy*
[3] *INFN Sezione di Firenze - Firenze, Italy*
[4] *Cell and Molecular Biology Department, Karolinska Institute - SE-17177 Stockholm, Sweden*

received 31 May 2006; accepted in final form 9 November 2006
published online 29 December 2006

PACS 87.23.Kg – Dynamics of evolution
PACS 87.17.Aa – Theory and modeling; computer simulation
PACS 82.20.-w – Chemical kinetics and dynamics

Abstract – The Chemoton model constitutes a minimalistic description of a protocell unit. The original formulation assumes three coupled chemical networks, representing a proto-metabolism, a template duplication and the membrane growth. An improved version is here proposed that explicitly incorporates the effects of the volume changes, due to the membrane growth. A stochastic mechanism is also introduced that mimics a stochastic source of error in the template duplication process. Numerical simulations are performed to monitor the time evolution of a family of protocells, under the chemoton hypothesis. An open-ended Darwinian evolution under the pressure of the environment is reproduced thus allowing to conclude that differentiation into species is an emergent property of the model.

Introduction. – Almost all life forms are composed by cells, fundamental constituting units which are able to self-replicate and evolve through changes in genetic information. These highly sophisticated devices are the product of about four-billion-year evolution, and, in this respect, they represent the relic of primordial life bricks, the protocells. The latter were most probably exhibiting only few simplified functionalities, that required a primitive embodiment structure, a protometabolism and a rudimentary genetics, so to guarantee that offsprings were "similar" to their parents [1–3].

Artificial protocells have not yet been reproduced in laboratory and intense research programs[1] are being established aiming at developing reference models [4,5] that capture the essence of the first protocells appeared on Earth and enable to monitor their subsequent evolution.

In 1971, Gánti [6] elaborated a pioneering theory that provides a minimalistic description of a protocell, termed Chemoton. Within this framework, growing and multiplying microspheres are controlled by a template duplication process. A membrane protects the inner bulk, while filtering the access of high energetic food, which is made available in the surrounding environment. The food is processed through chemical reactions and transformed into basic materials that are needed to stimulate both template duplication and membrane growth (see table I and table II). In the simplest scheme, the template polymer is assembled from one specific type of monomer and acts as an effective carrier of information: in fact its length determines a phenotype property of paramount importance, the division time. After such a time, the protocell attains a critical size, above which a division into two perfect halves occurs, we thus assume here that cell division would simply be due to a physically based process [4]. We also assume that each daughter contains an identical portion of material, which is equally shared from the mother constituents.

With respect to the original Chemoton picture, we here propose an improved formulation. First, let us notice that chemical reactions occur in a varying volume, which dynamically adjusts in time, as a result of the protocell membrane growth. This effect is not negligible and it is here explicitly accounted for[2].

[1] For instance, PACE —Programmable Artificial Cell Evolution— a European Integrated Project in the EU FP6–IST–FET Complex Systems Initiative.

[2] A forthcoming paper will be devoted to a complete analysis of the model, aiming in particular to investigate the dependence of the division time on the involved parameters [7]. The interested reader is also invited to consult the on-line supplementary material posted at http://www.fundp.ac.be/pdf/publications/58350.pdf [8].

Table I: *Chemoton hypothesis.* The chemical reactions describing the three coupled chemical subsystems. In the first column the autocatalytic cycle $A_1 \to 2A_1$ is reported, the second one focuses on the the membrane growth and the latter accounts for the template duplication. Here pV_{2N} denotes the double-stranded template constituted of $2N$ monomers V', whereas $(pV_{2N}pV_i)$, $i = 1, \ldots, 2N-1$, denotes a partially duplicated polymer where i-monomers have been copied. Direct reaction constants, k_j, are larger than the inverse ones, k'_j. We also assume the so-called strong replication condition [13] to hold true: k_6, k'_6 and k_7 are zero if the concentration of V' is below a given threshold V^*. The kinetic constants are set equal to the original Gánti's values, see also [8].

Metabolism	Membrane	Template
$A_1 + \bar{X} \underset{k'_1}{\overset{k_1}{\rightleftharpoons}} A_2$	$T' \xrightarrow{k_8} T^*$	$pV_{2N} + V' \underset{k'_6}{\overset{k_6}{\rightleftharpoons}} (pV_{2N}pV_1) + R$
$A_2 \underset{k'_2}{\overset{k_2}{\rightleftharpoons}} A_3 + \bar{Y}$	$T^* + R \underset{k'_9}{\overset{k_9}{\rightleftharpoons}} T$	$(pV_{2N}pV_i) + V' \xrightarrow{k_7} (pV_{2N}pV_{i+1}) + R$
$A_3 \underset{k'_3}{\overset{k_3}{\rightleftharpoons}} A_4 + V'$		$(pV_{2N}pV_{2N}) \longrightarrow pV_{2N} + pV_{2N}$
$A_4 \underset{k'_4}{\overset{k_4}{\rightleftharpoons}} A_5 + T'$		
$A_5 \underset{k'_5}{\overset{k_5}{\rightleftharpoons}} A_1 + A_1$		

Moreover, the above deterministic framework is further modified to accommodate for a simplified mutation mechanism: during the template duplication extra monomers are randomly added (removed) according to a pre–assigned probability, to mimic a stochastic source of errors in the reproduction process (see also [7,8]).

Within this novel theoretical scenario, we characterize the evolution of a population of protocells under the pressure of the environment, here synthesized by the amount of available food. In particular, we shall demonstrate that *differentiation into mutants' families*, hereafter termed *species*, is an emergent property of our model. This finding points to the fact that the Chemoton model represents not only a basic living unit, but also a fundamental *unit of evolution* [9]. We utilize here a definition of species for asexual beings based on differences between their genomes [10–12], which in our setting is equivalent to different lengths of the double-stranded templates.

Finally, it should be stressed that this remarkable feature cannot be observed for the original, *spherical-shaped*, Chemoton model [6,13] when both volume dependence and stochastic mutation are accounted for.

The model. – The original Chemoton model is schematically outlined in table I, where the chemical reactions involved are reported. The metabolic autocatalytic cycle is composed of five elementary steps, each involving one of the chemicals $A_1, \ldots, A_5$ and transforms the external nutrient (food), denoted by X, into internal material necessary to support the growth and self-reproduction processes. The food is here supposed to be buffered into the surrounding environment and therefore assumed constant. Products of the metabolism are the template and membrane precursors, termed respectively V' and T^*. The waste products Y are progressively eliminated so to maintain their level constant inside the membrane. The direct and inverse rate constants are, respectively, labeled k_i and k'_i.

The kinetic differential equations describing the temporal evolution of the concentrations of the chemical system are enclosed in table II. Here, the upper dot stands for time derivative and capital letters refer to the concentrations.

As anticipated, the above relations have to be modified to account for the fact that the volume changes in time and consequently the concentrations are dynamically rescaled. This observation translates in the following modification of the kinetic differential equations reported in table II. Consider a concentration c_i, at fixed volume, and assume the associate rate of variation to be specified by the function $f_i(\bar{c})$, intrinsically dependent on the other concentrations $\bar{c} = (c_1, \ldots, c_k)$. Then, if we allow the volume to change in time, the variation rate reads:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = f_i(\bar{c}) - c_i \frac{1}{V(t)} \frac{\mathrm{d}V(t)}{\mathrm{d}t}, \tag{1}$$

where $V(t)$ stands for the protocell volume at time t.

Note however that the kinetic formulation here discussed needs to be self-consistently completed to accommodate for the time evolution of the volume, which in turns amount to specify a given shape of the membrane, the surface growth being successfully captured by the model.

Table II: *Kinetic differential equations.* From the chemical reactions reported in table I, one deduces the above differential equations describing the evolution in time of the concentrations of the involved chemicals. To lighten notations, we denote the concentrations with capital letters. The last equation of the middle column describes the growth of the surface size as a result of the progressive attachment of the membrane molecules [14]. All equations, except the one for $\dot{S}$, have to be modified according to eq. (1).

$\dot{A}_1 = 2\left(k_5 A_5 - k_5' A_1^2\right) - k_1 A_1 \bar{X} + k_1' A_2$	$\dot{T}' = k_4 A_4 - k_4' A_5 T' - k_8 T'$	$\dot{pV_{2N}} = 2k_7(pV_{2N}pV_{2n-1})V'$ $+k_6'(pV_{2N}pV_1)R - k_6 pV_{2N}V'$
$\dot{A}_2 = k_1 A_1 \bar{X} - k_1' A_2 - k_2 A_2 + k_2' A_3 \bar{Y}$	$\dot{T}^* = k_8 T' - k_9 T^* R + k_9' T$	$\dot{(pV_{2N}pV_1)} = k_6 pV_{2N}V' - k_6'(pV_{2N}pV_1)R$ $-k_7(pV_{2N}pV_1)V'$
$\dot{A}_3 = k_2 A_2 - k_2' A_3 \bar{Y} - k_3 A_3 + k_3' A_4 V'$	$\dot{T} = k_9 T^* R - k_9' T - k_{10} T S$	$\dot{(pV_{2N}pV_{i+1})} = k_7(pV_{2N}pV_i)V'$ $-k_7(pV_{2N}pV_{i+1})V'$
$\dot{A}_4 = k_3 A_3 - k_3' A_4 V' - k_4 A_4 + k_4' A_5 T'$	$\dot{S} = k_{10} T S$	$\dot{V}' = k_3 A_3 - k_3' A_4 V' + k_6'(pV_{2N}pV_1)R$ $-k_6 pV_{2N}V' - k_7 \sum_{i=1}^{2N-1}(pV_{2N}pV_i)V'$
$\dot{A}_5 = k_4 A_4 - k_4' A_5 T' - k_5 A_5 + k_5' A_1^2$		$\dot{R} = k_6 pV_{2N}V' - k_6'(pV_{2N}pV_1)R + k_9' T$ $-k_9 T^* R + k_7 \sum_{i=1}^{2N-1}(pV_{2N}pV_i)V'$

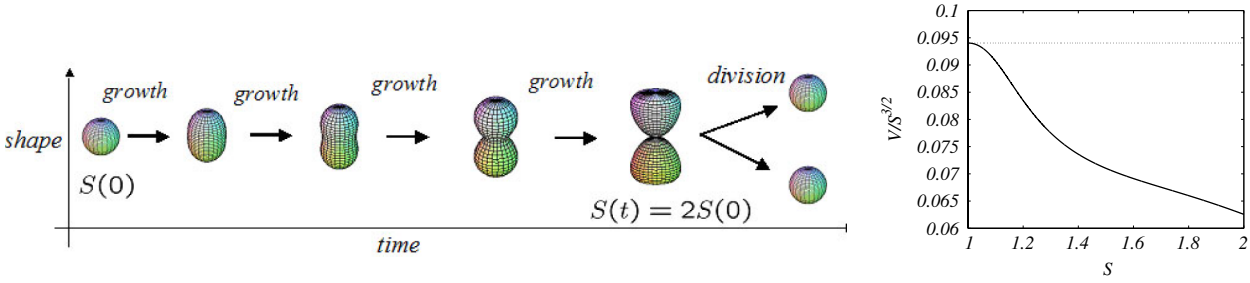

Fig. 1: Left panel: *time snapshots of the protocell evolution.* At $t = 0$ the protocell is assimilated to a sphere whose surface is $S(0) = 1$. At the division time, the surface has doubled, $S(t) = 2S(0)$ and the shape resembles an hourglass. Right panel: *the ratio $V/S^{3/2}$ against the surface size, S.* The black solid line stands for our model, while the dotted line refers to the case of a sphere. Note that, by assumption, at $S = 1$ the two curves coincide.

To this end we observe that cells can divide through a pinch region which makes them schematically resemble to an *hourglass*. We therefore impose the protocell to progressively deform its initial spherical shape and eventually assume the characteristic pinched profile when the duplication initiates. This basic mechanism is outlined in the left panel of fig. 1.

Deviations from the ideal spherical distribution play a central role. For a given surface, in fact, the shape determines the volume enclosed by the membrane and consequently affects the relative concentrations. The sphere maximizes the inner volume which in turn implies that other, more realistic shapes would delimit smaller portions of space, hence resulting in higher concentration. This observation is exemplified in the right panel of fig. 1 where the ratio $V(t)/[S(t)]^{3/2}$ is displayed as a function of the surface size for the case considered in this study. Observe that, for the case of a sphere such a ratio, is constant and equal to $1/(6\sqrt{\pi})$. Finally, let us stress that having introduced a smooth splitting mechanism enables to significantly reduce the functional discontinuities observed within the original Chemoton picture [13] in conjunction with a duplication event. A more detailed discussion can be found in [8].

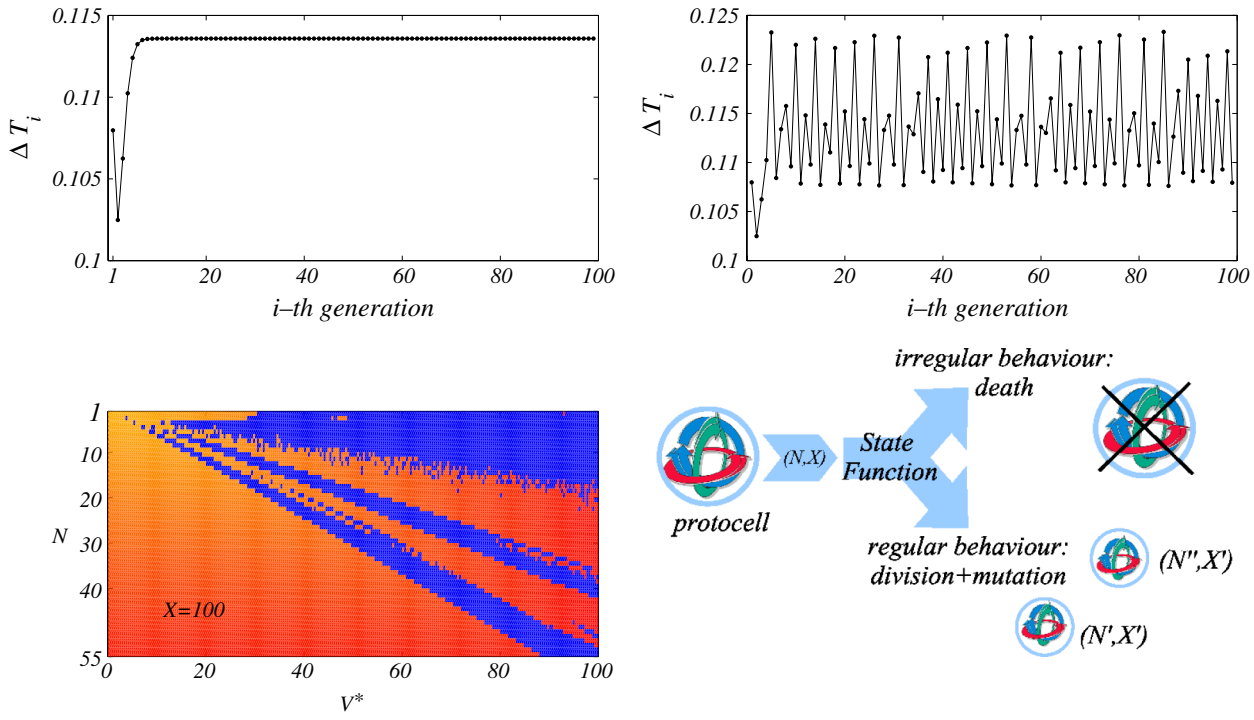

Fig. 2: Upper panels: *the division time*. We show the division time as a function of the generation number, for a regular (left panel), *i.e.* periodic, behavior. After a short transient, the division time attains a fixed value: protocells belonging to different generations employ the same amount of time to take the division step to completion. On the right panel an irregular, *i.e.* non-periodic, behavior. The division time depends on the generation number and it never achieves an asymptotic constant value. Lower panels: *the state function*. Left: a graphical representation of the state function for polymers composed by 1 up to 55 monomers whose polymerization ability ranges from 0 (pronounced ability to self-replicate) to 100 (reduced capacity to self-replicate). Colors (on-line) are assigned according to the classification in regular and irregular behavior: blue dots refer to irregular sequences, whereas regular orbits are represented in red, the faster the division period the lighter the color-coding. A fixed amount of available food is assumed. Right: schematic use of the state function. N denotes the polymer length while X the available food. Protocells exhibiting non-periodic behavior will eventually die, while each protocell characterized by a regular behavior divides into two new offsprings whose polymer lengths can be different from the mother ones because of the mutations.

In the following we shall adopt our improved theoretical framework to numerically investigate the time evolution of a family of protocells. Interestingly, no complete stochastic simulations of a Chemoton-like model have been tempted so far, the reason being ascribed to the huge computations involved if we would naively track the evolution of each single protocell while the population size increases exponentially due to the combined effect of the duplication and mutation mechanisms. We here overcome these technical difficulties by proposing an alternative numerical strategy, shortly introduced in the forthcoming section.

Results and conclusions. – Within this novel scenario, the time evolution of a *single protocell* is monitored as a function of a number of selected variables of key relevance, *e.g.* the amount of food initially available and the template length at start. This knowledge translates into a state function[3] which unambiguously predicts the ultimate fate of the protocell, once the initial conditions are specified, *e.g.* the length of the polymer (N) and its ability[4] to self-replicate (V^*). More precisely, the division time is numerically calculated for a large number of parameters values, (N, V^*): periodic and non-periodic orbits are consequently identified (see the upper panels in fig. 2) and this knowledge enables to partition the parameters space into basins of regular or irregular (chaotic) dynamics (see the lower panels in fig. 2 and [8]). Observe that these two opposite behaviors are not artifacts but instead represent an intrinsic pecularity of the model: they are in fact the signatures of the presence or absence of synchronization between the three subnetworks.

Assume now that the external food is periodically available, each cycle being for instance associated to tides in the pond were the protocells supposedly live, or the metabolic process to be driven by a periodic source of energy, *e.g.* the light from the sun. Then the regular functioning of the model, *i.e.* periodic growth and division

[3]The latter corresponds to using a cache for the fitness function.

[4]Let us recall that V^* has been previously defined as the threshold below which the polymerization cannot be performed, but this can also be interpreted as the ability of the double-stranded template to replicate.

in time of the protocell, relies on the synchronization of the three chemical subsystems. Motivated by the above synchronization ansatz and due to the absence of intermittency in the protocell dynamics, *i.e.* regular behaviors which became irregular —or vice versa— waiting long enough, we here hypothesize that only protocells with regular behavior give rise to next-generation offsprings, while non-synchronized protocells will eventually die[5]. As already remarked, offsprings from a protocell that displays an irregular behavior could present, after the division event, lower concentration of double-stranded template. Hence, we decide to impose a strong selection mechanism by assuming that only the protocells that optimize the resources, food *vs.* duplication time, can eventually survive.

The protocells are subjected to the pressure of the environment, here exemplified through the amount of available food, which changes in time and depends self-consistently on the size of the population: the larger the number of protocells, the lower the amount of available food, and vice versa.

Using the state function concept, computed for different values of food X, we are able to perform for the first time simulations on a large population of protocells, following their evolution for long intervals of time. Each protocell behavior can be described by more than one hundred of equations. Computing the state function once for all, enables to formally assimilate each protocell belonging to the population under scrutiny to a black box, which produces specific output, once selected input values are provided (see the lower right panel of fig. 2 for a schematic representation of the use of the state function; we also refer the interested reader to [8]). This reduces considerably the computational costs and allows to significantly enhance the statistics over previous investigations.

In conclusion, as clearly displayed in fig. 3, the simulations mimic an effective Darwinian evolution of a population of protocells, obeying to the Chemoton hypothesis, and being subjected to a mutation mechanism and to the pressure of the environment. Figure 3 refers to one typical run: such result is however quite general and robust to small change of the parameters involved.

The main result of our analysis is that the speciation is an emergent property of the model[6]: protocells with different templates can develop from a common ancestor, while competiting for available food. The occurrence of a speciation mechanism in a Chemoton-like population is here demonstrated for the first time and contributes to shed new light into the important issue of protocells

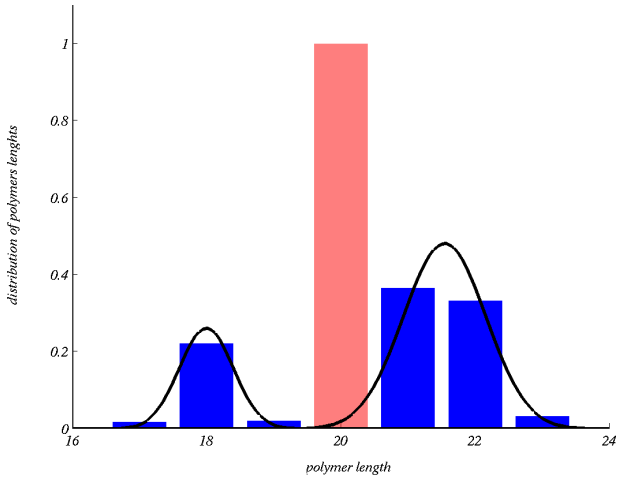

Fig. 3: (Color on-line) Distribution of polymers length after 20 generations (blue), starting with 5 identical protocells with polymer length of 20 monomers (red). Two species are clearly identified, resulting from the selection of the fittest mutated protocells, in the competition for food. The solid line interpolates the result of the numerical experiments. This picture represents a generic evolution: distinct runs relative to different choices of the parameters result in similar behaviors.

evolution and dynamics: the Chemoton hypothesis defines not only a reliable unit of life, but also a unit of evolution. Importantly, this scenario does not apply to the original Chemoton description, the geometrical aspects related to the process of division, and incorporated in our model, playing a role of paramount importance [8].

* * *

The work of the first author has been partially funded by PACE (Programmable Artificial Cell Evolution), a European Integrated Project in the EU FP6–IST–FET Complex Systems Initiative under contract FP6–002035, which is gratefully acknowledged. The numerical code is implemented in Matlab and can be delivered upon request.

REFERENCES

[1] Alberts B. *et al.*, *Molecular Biology of the Cell* (Garland, New York) 2002.
[2] Rasmussen S. *et al.*, *Science*, **303** (2004) 963.
[3] Szostak D., Bartel P. B. and Luisi P. L., *Nature*, **409** (2001) 387.
[4] Luisi P. L., Ferri F. and Stano P, *Naturwissenschaften*, **93** (2006) 1.
[5] Rasmussen S., Chen L., Stadler B. and Stadler P., *Origins Life Evol. Biosphere*, **34** (2004) 171.
[6] Gánti T., *Chemoton Theory: I) Theory of Fluid Machineries and II) Theory of Living Systems*, edited by Paul G. Mezey (Kluwer Academic/Plenum Publishers, New York) 2003.

[5]However, it should be mentioned that recent results [15] enables us to obtain an analogous speciation mechanism, while completely relaxing the above hypothesis.

[6]It is worth emphasising that the speciation mechanism here outlined is neither trivial nor *a priori* predictable. It is indeed resulting from a complex interplay between the postulated mutation scheme and the effect of the *resonances*, an intrinsic feauture of our modified Gánti's model that manifests itself as forbidden regions in the state function space.

[7] Carletti T., *Stability, mutations and evolution in a population of Chemotons,* submitted to *J. Theor. Biol.* (2006).
[8] Carletti T. and Fanelli D., On-line complementary material annexed to this paper and posted online at `http://www.fundp.ac.be/pdf/publications/58350.pdf` (2006).
[9] Szathmary E., Santos M. and Fernando C., *Top. Curr. Chem.*, **259** (2005) 167.
[10] Cohan F. M., *Annu. Rev. Microbiol.*, **56** (2002) 457.
[11] Konstantinidis T. K. and Tiedje J. M., *Proc. Natl. Acad. Sci. U.S.A.*, **102** (2005) 2567.
[12] Moreno E., *Rev. Biol. Trop.*, **50** (2002) 803.
[13] Munteanu A. and Solé R. V., *J. Theor. Biol.*, **240** (2006) 434.
[14] Singer J. and Nicolson G. L., *Science*, **175** (1972) 720.
[15] Carletti T. and Fanelli D., contribution to *BIOMAT2006 International Symposium, November 2006*, to be published.