Regular Article

Finite-size corrections to scaling behavior in sorted cell aggregates

A.V. Klopper[1,2], G. Krens[3,a], S.W. Grill[1,2,b], and C.-P. Heisenberg[3]

[1] Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstraße 108, D-01307 Dresden, Germany
[2] Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, D-01187 Dresden, Germany
[3] Institute of Science and Technology Austria, Am Campus 1, A-3400 Klosterneuburg, Austria

Abstract. Cell sorting is a widespread phenomenon pivotal to the early development of multicellular organisms. *In vitro* cell sorting studies have been instrumental in revealing the cellular properties driving this process. However, these studies have as yet been limited to two-dimensional analysis of three-dimensional cell sorting events. Here we describe a method to record the sorting of primary zebrafish ectoderm and mesoderm germ layer progenitor cells in three dimensions over time, and quantitatively analyze their sorting behavior using an order parameter related to heterotypic interface length. We investigate the cell population size dependence of sorted aggregates and find that the germ layer progenitor cells engulfed in the final configuration display a relationship between total interfacial length and system size according to a simple geometrical argument, subject to a finite-size effect.

1 Introduction

The segregation and sorting of cells with different fates are key processes in the development of all multicellular organisms. The sorting of cells has long been thought to result primarily from the disparate adhesiveness of the participating cells [1]. However, recent studies have provided evidence that differential intercellular adhesion is not sufficient to explain cell sorting and that other cellular properties, such as differences in actomyosin-dependent cell cortex tension, cell polarization and cell migration are also critical to this process [2]. Quantification of the relative contributions of these properties requires accurate recording and analysis of the sorting process in space and time. While extensive experimental studies have provided the basis for our current understanding, these investigations have been based on three-dimensional cell sorting events analyzed in two dimensions, owing to the difficulties associated with three-dimensional imaging and image analysis [3–8].

Simulations have been remarkably successful in capturing generic cell sorting behavior, despite being largely restricted to simplified model systems and two-dimensional geometries [9–11]. While the extrapolation to three dimensions is non-trivial [3,9–12], these investigations have proven particularly useful in the development of techniques for quantifying and analyzing sorting behavior. One such technique involves calculating the heterotypic interface length as the cells reorganize. This is trivial in the case of two-dimensional lattice simulations [9] but less straightforward for free particle simulations [11] and experimental data in three-dimensional geometries. The definition of an order parameter measuring local heterotypical density provides a convenient alternative which is related to the interface length in a segregated cell population [11]. However, simulation data examined via this technique yield unusual exponent values which suggest that the total interface length between cellular clusters in the segregated phase is not geometrically related to system size. This discrepancy poses a problem for the analysis of experimental studies, wherein one requires a means with which to compare diverse biochemical and genetic conditions for systems that may vary in size. Here, we investigate the applicability of the proposed order parameter to the study of primary embryonic cells during sorting.

We report on extensive experimental studies of sorting embryonic zebrafish germ layer progenitor cells in three dimensions. Our results indicate that the relationship between heterotypical interface length and system size is indeed geometrically driven. We construct a geometrical argument to describe this relationship and show that a finite-size effect is implicated in the modification of the expected exponents. The investigation presents a platform from which further biochemically motivated studies can be launched.

---

[a] A.V. Klopper and G. Krens share joint first authorship.
[b] e-mail: grill@mpi-cbg.de
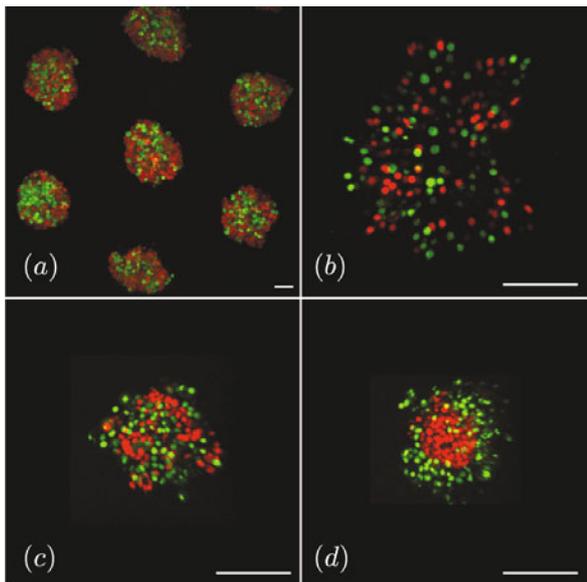
Fig. 1. (Color online) Imaging data for different time points in a single experiment. Scale bar $= 100\,\mu$m. (a) Micro-molds are used to isolate small populations of ecto- and mesoderm cell mixtures labeled fluorescently with red and green nuclei, respectively. (b) Initial images show homogeneously mixed cells distributed throughout the mold. (c) Cells aggregate together on a time scale of roughly 100 minutes. (d) Imaging after sorting clearly shows the segregation of the two cell populations.

2 Cell sorting order parameter

Fluorescent labeling of zebrafish nuclei facilitates the imaging of ectoderm and mesoderm germ layer progenitor cells *in vitro*, while they engage in the aggregation and sorting behavior observed in early development. Imaging this process with high accuracy in three dimensions is not trivial. Segregation occurs over long time periods during which the primary cells are removed from their native environment. Optimization of image sampling and label intensity reduces the exposure of the cells to phototoxic effects during this extended image acquisition. Furthermore, under normal experimental conditions (the so-called "hanging drop" preparation) aggregates are subject to drift and deformation during prolonged time-lapse imaging, complicating the collation of three-dimensional images. We implement an ideal solution to this problem by performing the experiments in micro-molds [13] with non-adhesive and non-deformable surfaces.

For this study, we inject mRNAs into the embryos at the single-cell stage, inducing either ectoderm or mesoderm cell fates and labeling nuclei in red and green respectively (see appendix A for details). We then isolate a mixture of the two cell types, in approximately equal quantities, and deposit them in micro-molds as shown in fig. 1(a). The subsequent demixing process is imaged every three minutes in three dimensions over a five-hour period. Only nuclei within a hemispherical section of the micromold are imaged. Snapshots of the acquired image data are illustrated in figs. 1(b)-(d). Nuclei positions are de-

termined at each time point by means of an automated template-matching algorithm [14].

After a gravitationally driven settling into the microwells, the cells undergo aggregation as they segregate, resulting in the green-labeled mesoderm cells engulfing the red-labeled ectoderm cells (see fig. 1(d)). The cells also divide during imaging and population sizes initially increase exponentially as shown for the same experiment in fig. 2(a). In this respect, our analysis is significantly more challenging than that of simulated data, in which cell populations are maintained at a constant size for the duration of the process. Nevertheless, we note that the conditions under which cells segregate seem to provide for stabilization of their number and we are able to estimate sorted population sizes based on the saturating values of exponential functions fitting the curves in fig. 2(a).

We define a local order parameter for each cell type,

$$\gamma_A(t) = \frac{1}{N_A} \sum_{i=1}^{N_A} \frac{n_B^i}{n_A^i + n_B^i}\,, \qquad (1)$$

where subscripts A and B denote either mesoderm or ectoderm germ layer progenitor cell populations and N_A refers to the total number of cells in the population of interest [11]. We define n_A^i and n_B^i as the respective numbers of similar and dissimilar cells neighboring cell $i \in A$ so that the order parameter forms a population average of the fraction of dissimilar cells neighboring a given cell species. For a homogeneous distribution of two different cell types, the order parameter is simply determined by the ratio of one population size to the other. In the segregated phase, it is proportional to the interface length between clusters [11].

From the nuclei position data, we determine the nearest neighbors of each imaged nucleus via a customized radial cutoff which ensures that the average number of nearest neighbors $\langle n \rangle$ across the population approaches that of close packing in 3D *i.e.*, $\langle n \rangle = 12$. This is achieved by calculating pairwise distances for the entire aggregate and selecting a radial cutoff larger than the k-th smallest distance computed, where $k = \langle n \rangle (N_{\mathrm{ecto}} + N_{\mathrm{meso}})/2$. The cutoff decreases with time as the cells aggregate and may be fit with an exponential function as shown in the example given in fig. 2(b). This tendency reflects the aggregation of the cell population. We note that the behavior of the order parameter is not critically dependent on the choice of $\langle n \rangle$, given the bounds of error associated with our analysis. We have confirmed this by comparing our results with those obtained when $\langle n \rangle \in \{8, 9, \ldots, 14\}$ (data not shown). The broken line (also shown in (a) and (c)) indicates $t = 2\tau$ where τ is the decay constant associated with the fit in fig. 2(b), signifying an end to the aggregation process.

In fig. 2(c) we plot the order parameter in eq. (1) for both ectoderm and mesoderm cell populations for the experiment in fig. 1 as a function of time. A decrease in the order parameter coincides with an increase of the number of homotypic neighbors on average across the population. The behavior of the order parameter is well described by
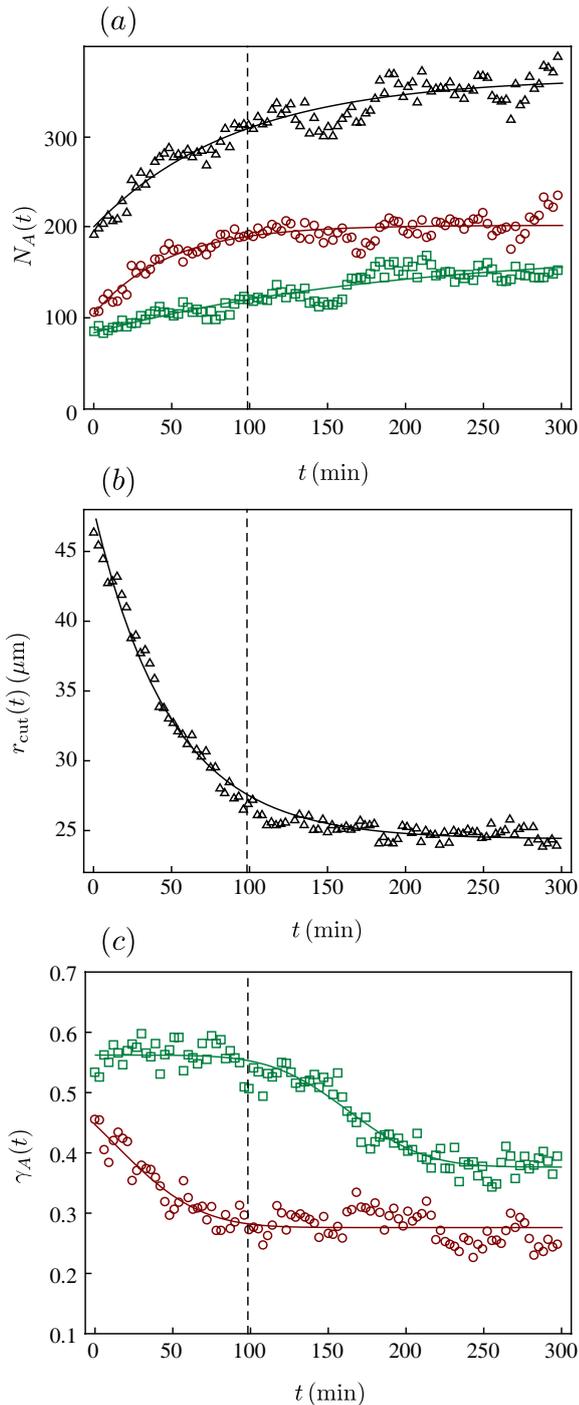
Fig. 2. (Color online) Data analysis for the experiment depicted in fig. 1. Black triangles denote the entire population of cells; red circles denote ectoderm (engulfed) cells; green squares denote mesoderm (engulfing) cells. (a) $N_A(t)$ describes cell population numbers, which appear to saturate over time. Subscript A refers to ecto- or mesoderm cells and fit lines are exponential functions. (b) Radial cutoff $r_{\mathrm{cut}}(t)$ used to locate cell nearest neighbors. The solid line is an exponential fit and the broken line (also shown in (a) and (c)) indicates $t = 2\tau$, where τ is the decay constant associated with the fit. (c) $\gamma_A(t)$ is the order parameter used to quantify sorting. It shows a saturating tendency in the long time limit, described well by error function fits.

an error function ($\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{d}\lambda\, e^{-\lambda^2}$) for both cell types. In the example shown, the order parameter for the ectoderm cells decreases immediately after seeding, which is reflected in the observation of early clustering of the red labeled cells in fig. 1. The fact that this clustering does not affect the order parameter for the mesoderm cells (which remains constant in the first 100 minutes) can be attributed to the difference in the rates at which the two populations increase, illustrated in fig. 2(a).

We expect the saturating value of the order parameter to depend on system size in an essentially geometrical manner, given the behavior of the cells in the segregated phase in fig. 1(d). The nature of this relationship may be determined by first inferring the number of cells of each kind in the long time limit. Despite fluctuating cell numbers, we are able to estimate relative population sizes from the saturating values of exponential functions fitting the curves in fig. 2(a). The raw data from 48 experiments are fit in the manner exemplified by figs. 2(a) and (c). By considering only data from experiments which are not compromised by imaging problems, we are able to assemble a robust data set ($R^2 > 0.95$) of 31 experiments with which we obtain saturating values $\gamma_A^\star \equiv \gamma_A(t \to \infty)$ and $N_A^\star \equiv N_A(t \to \infty)$ for both ecto- and mesoderm cell populations. Of the experiments that do not meet our imaging quality criteria, we note that the majority constitute attempts to acquire data for very large population sizes. Clearly, the difficulties associated with imaging large system sizes set an upper limit on the parameter range to which our study has access.

3 Size dependence of segregation

In the segregated configuration, the sum in eq. (1) comprises contributions from interfacial cells only and one thus expects its value to scale with system size in the same manner as does the ratio of interfacial to bulk cells. This implies that

$$\gamma_A^\star \sim N_A^{\star\,\mu_A}, \qquad (2)$$

where $\mu_A < 0$ such that the order parameter for an infinitely large system approaches $\gamma_A^\star = 0$. Since most of the mesoderm cells observed in our experiments form a relatively thin engulfing layer (see fig. 1(d)), the order parameter should display a weak dependence on system size corresponding to small μ_{meso}. Conversely, the order parameter for the engulfed ectoderm cells is expected to scale with population size according to the spherical surface-to-volume ratio with $\mu_{\mathrm{ecto}} = -1/3$. However, within the range of system sizes probed by our experiments, a finite-size effect accounts for the observation of a modified exponent. This is manifested in a surface-to-volume ratio measured in units of cellular size: in the lower limit corresponding to a solitary ectoderm cell, this ratio is unity; in the limit of large population sizes, the exponent approaches $\mu_{\mathrm{ecto}} = -1/3$.

In fig. 3 we plot the saturating values of $\gamma_A^\star$ and $N_A^\star$ obtained from our experiments. The data are best fit with exponents $\mu_{\mathrm{ecto}} = -0.24 \pm 0.12$ for ectoderm cells and
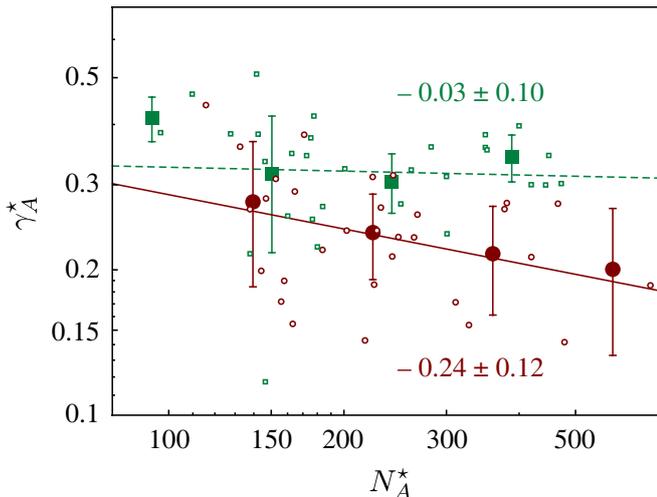
Fig. 3. (Color online) The saturating value for the order parameter $\gamma_A^\star$ as a function of the size of the population $N_A^\star$ for the ectoderm (red circles) and mesoderm (green squares) cells. Empty markers represent single experiment values which are fit with solid (ectoderm) and dashed (mesoderm) lines with slopes corresponding to inset exponents. Filled markers represent binned values and error bars indicate the standard deviation of the data in each bin.

Fig. 4. (Color online) The saturating value of the order parameter for the engulfed ectoderm cells $\gamma_{\mathrm{ecto}}^\star$ depends on the size of the population $N_{\mathrm{ecto}}^\star$ in the manner predicted by eq. (4), plotted as a dashed line. The dotted line has a slope of $-1/3$, indicating the surface-to-volume ratio expected in the large population limit for comparison. Data are binned and error bars indicate the standard deviation of the data in each bin. The inset shows an enlargement of the area of interest.

$\mu_{\mathrm{meso}} = -0.03 \pm 0.10$ for mesoderm cells. These fits also appear in fig. 3, along with binned data points which are best fit with exponents $\mu_{\mathrm{ecto}} = -0.26$ and $\mu_{\mathrm{meso}} = -0.02$. These exponents suggest that the fraction of mesoderm cells lying at the interface scales weakly with system size and also suggests that the scaling behavior for the ectoderm cells deviates from the large population limit.
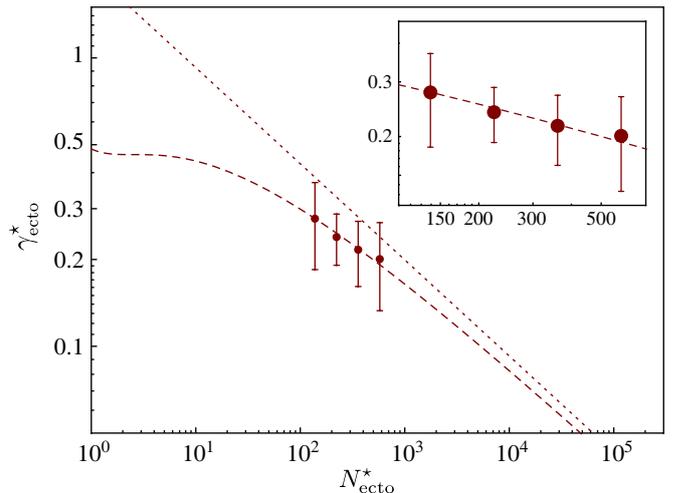
We can understand this behavior by computing the ratio of interfacial to bulk ectoderm cells in terms of a randomly packed, approximately spherical ensemble of cells. We estimate the inferred volume occupied by the segregated ectoderm cell population, V_{ecto}, by

$$\eta \cdot V_{\mathrm{ecto}} \approx N_{\mathrm{ecto}}^\star \cdot v_{\mathrm{ecto}}, \tag{3}$$

where η is the packing fraction of the cells and v_{ecto} is the spherical cellular volume. By choosing this parameter to reflect the likely value for an ensemble of spheres centered at the imaged nuclei, we are able to estimate the number of ectoderm cells neighboring mesoderm cells within the hemispherical section. These cells are expected to lie within an outer shell of the half-sphere of thickness equivalent to the diameter of a single cellular sphere such that the expected saturating value for the order parameter scales as

$$\gamma_{\mathrm{ecto}}^\star \sim \left(1 - 2\left(\frac{\eta}{2N_{\mathrm{ecto}}^\star}\right)^{1/3}\right)^3. \tag{4}$$

This expression approaches $(N_{\mathrm{ecto}}^\star)^{-1/3}$ for very large $N_{\mathrm{ecto}}^\star$, but this limit is well outside the range of our imaged

data. In fig. 4 we plot eq. (4) with $\eta = 0.64$ (consistent with random packing) on a log-log scale (dashed line), along with $(N_{\mathrm{ecto}}^\star)^{-1/3}$ (dotted line). The binned ectoderm data shown in fig. 3 are also reported here. While the range of systems to which our assay has access is not large enough to confirm the scaling behavior proposed in eq. (4), the description is clearly consistent with our data. In addition to the difficulties associated with imaging large aggregates, the major obstacle with respect to imaging smaller cell populations rests with the inability to induce engulfment while maintaining a roughly equal ecto- to mesoderm cell ratio. Nevertheless, in all experiments analyzed we observe complete engulfment of the ectoderm cells by mesoderm cells. The lower limit on the number of mesoderm cells required for engulfment is itself subject to non-trivial dependence on the size of the ectoderm cell population. The inset plot in fig. 4 further illustrates the close agreement between data and the expected scaling behavior for accessible population sizes. This agreement highlights the importance of considering the reported finite-size effect when analyzing experimental data and illustrates that scaling exponents from estimates on large systems are not necessarily reproduced in experiment.

The geometric argument presented in eq. (4) can be readily modified to describe scaling behavior in two dimensions. However, the finite-size correction to the expected exponent in the two-dimensional case is only significant in systems smaller than those typically analyzed. We would thus expect such systems (be they *in vitro* or *in silico*) to display the simple geometric scaling associated with very large systems. Note that this prediction is at odds with the unusual exponent reported in ref. [11].

4 Conclusion

Ongoing investigations into cellular reorganization during embryogenesis are heavily reliant on developing an understanding of sorting processes in three dimensions over time. Simulations and static imaging have been instrumental in revealing the most likely cell properties involved in these processes, but results have been inconclusive and largely restricted to two dimensions. Here we investigate cell sorting phenomena dynamically in three dimensions by assembling aggregates of different types of primary embryonic zebrafish cells. We use a novel method for real time observation and analysis of the demixing of cells during cell sorting *in vitro*. The results reported here show a size-dependent segregated phase and our analysis, using a previously formulated order parameter, presents a robust means of understanding this behavior. We find that the germ layer progenitor cells engulfed in the final configuration relate the total interfacial length to system size in terms of a simple geometrical argument, subject to a finite-size effect. A related argument holds for the engulfing cells, which gives rise to weak size dependence of the order parameter, valid for the system sizes and the 1 : 1 cell ratio used in our experiments.

Importantly, the reported finite-size correction is dependent on the dimensionality of the system. By recording and analyzing three-dimensional cell sorting events in three dimensions, we circumvent the difficulties this effect poses for the interpretation of two-dimensional analysis. Our study presents an ideal starting point for further research into the driving principles of cellular aggregation and sorting. Known players such as cell migration, adhesion and cortical tension are all sensitive to biochemical perturbation. We are currently investigating the effect of these cellular properties on cell sorting using the techniques developed for this study.

Appendix A. Materials and methods

Ectoderm and mesoderm cells were obtained as described in ref. [2]. Wild-type tub longfin embryos were injected at the single cell stage with *lefty1* mRNA (100 pg) or *cyclops* mRNA (100 pg) and *casanova* morpholino oligonucleotides (2 ng; GeneTools). Fluorescent labeling was achieved by co-injecting mRNA, encoding either histone2A-mCherry or histone2B-EGFP (100 pg), resulting in red- and green-labeled nuclei, respectively.

Embryos were kept at 31 °C until dissociation into single cells at sphere-stage (4 hpf), as described in ref. [15]. After blastoderm dissociation, the two differentially labeled germ layer progenitor cell types (ectoderm and mesoderm) were mixed in a 1 : 1 ratio. Of this mixture, containing 5×10^5 cells/mL, 35 μL was seeded in a micro-mold comprising 56 wells. The cells were then left to settle, aggregate and sort in the individual wells. In order to optimize imaging, cell demixing was recorded in 3 of the 56 wells for at least five hours in three dimensions over time at 31 °C, by acquisition of 5 μm-spaced z-stacks of the aggregate in two channels every three minutes, using a Leica SP5 confocal microscope equipped with a 20× dipping lens.

Micro-molds were made from a PDMS negative [13] containing 56 pegs corresponding to the 56 wells, each with a diameter of $\pm 200 \, \mu$m and a height of $\pm 800 \, \mu$m, by pouring liquid 2% Ultrapure© Agarose (Invitrogen) in dH$_2$O and allowing it to gel on ice. The agarose micro-mold gel was removed from the PDMS negative and equilibrated for two hours with 2 mL of CO$_2$-independent DMEM /F12 tissue culture medium (Invitrogen) at 31 °C. Nuclei positions were determined at each time point by means of an automated template-matching algorithm [14].

References

1. M.S. Steinberg, Science **141**, 401 (1963).
2. M. Krieg, Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D.J. Müller, C.-P. Heisenberg, Nat. Cell. Biol. **10**, 429 (2008).
3. J.C.M. Mombach, J.A. Glazier, R.C. Raphael, M. Zajac, Phys. Rev. Lett. **75**, 2244 (1995).
4. D.A. Beysens, G. Forgacs, J.A. Glazier, Proc. Natl. Acad. Sci. USA. **97**, 9467 (2000).
5. D. Duguay, R.A. Foty, M.S. Steinberg, Dev. Biol. **253**, 309 (2003).
6. E.E. Robinson, K.M. Zazzali, S.A. Corbett, R.A. Foty, J. Cell Sci. **116**, 377 (2003).
7. J.C.M. Mombach, D. Robert, F. Graner, G. Gillet, G.L. Thomas, M. Idiart, J.-P. Rieu, Physica A **352**, 525 (2005).
8. R.A. Foty, M.S. Steinberg, Dev. Biol. **278**, 255 (2005).
9. F. Graner, J.A. Glazier, Phys. Rev. Lett. **69**, 2013 (1992).
10. G.W. Brodland, J. Biomech. Eng. **124**, 188 (2002).
11. J.M. Belmonte, G.L. Thomas, L.G. Brunnet, R.M.C. de Almeida, H. Chaté, Phys. Rev. Lett. **100**, 248702 (2008).
12. M.S. Hutson, G.W. Brodland, J. Yang, D. Viens, Phys. Rev. Lett. **101**, 148105 (2008).
13. A.P. Napolitano, D.M. Dean, A.J. Man, J. Youssef, D.N. Ho, A.P. Rago, M.P. Lech, J.R. Morgan, Biotechniques **43**, 494 (2007).
14. J. Stühmer, Thesis, Technische Universität Dresden (2007).
15. E.M. Schötz, R.D. Burdine, F. Jülicher, M.S. Steinberg, C.-P. Heisenberg, R.A. Foty, HFSP J. **2**, 42 (2008).