Collective regulation of the amoeboid motility : the role of short and long-range interactions in vegetative Dictyostelium discoideum

Joseph d'Alessandro

Université Claude Bernard - Lyon 1
Institut Lumière Matière

THÈSE
Pour l'obtention du
Diplôme de Doctorat
Spécialité : Physique

Collective regulation of the amœboid motility: the role of short and long-range interactions in vegetative *Dictyostelium discoideum*

Travail présenté par

Joseph d'Alessandro

Soutenance prévue le 16 mars 2016 devant le jury composé de :

M. Christophe ANJARD	Invité
M. Éric CLÉMENT	Examinateur
M. Jean-Marc DI MEGLIO	Rapporteur
M. Yoshinori HAYAKAWA	Invité
M. Clément NIZAK	Examinateur
M. Jean-Paul RIEU	Directeur de thèse
Mme Charlotte RIVIÈRE	Codirectrice de thèse
M. Raphaël VOITURIEZ	Rapporteur

Institut Lumière Matière
UMR5306 CNRS
Université Claude Bernard Lyon 1
Domaine Scientifique de La Doua
Bâtiment Kastler, 10 rue Ada Byron
69622 Villeurbanne CEDEX, France

Remerciements

Comme il est de coutume, après rédaction,
Soutenance, et puis pot, passée l'inaction
Nécessaire au repos si longtemps envié,
Le docteur reprend plume, ou plutôt clavier.

Les quatre lignes qui précèdent, vestiges de l'ambition aussi démesurée que vite ravalée d'écrire le présent chapitre en alexandrins, témoignent en quelque sorte de ce que peut être une thèse : une suite de projets tous trop ambitieux, une montagne que le thésard impétueux tente d'attaquer par sa face la plus raide et la plus lisse, attitude qui ne peut mener qu'à l'impasse, ou à la chute. Pourtant, contrairement à ces vers trop vite délaissés, mes quelques années lyonnaises m'auront permis de gravir patiemment les pentes du doctorat et d'ouvrir à ma vue de nouveaux horizons. Ce chemin parcouru, je le dois en grande partie aux nombreuses personnes qui l'ont jalonné, qui en me donnant un coup de pouce dans un raidillon, qui en m'orientant aux carrefours, ou même parfois par leur simple présence, m'évitant l'ennui du marcheur solitaire. Un grand merci à tous, et mes excuses par avance au lecteur qui se sentira lésé par la liste suivante – il y en aura forcément, mais que cet oubli ne soit pas interprété comme un manque de reconnaissance !

Pour faire les choses dans le désordre, commençons par la fin. Merci donc à tous les membres du jury pour le temps que vous avez consacré à évaluer mon travail. Merci surtout pour ces discussions et ces remarques nombreuses qui m'ont permis d'enrichir la vision que j'ai de mon sujet. Merci aussi de m'avoir permis en parallèle de travailler mon endurance, au prix pour certains d'entre vous d'un pot raccourci. Au delà du jour de la soutenance, je suis très reconnaissant à Clément Nizak pour les interactions régulières, toujours constructives et valorisantes, que nous avons eu durant ces quelques années. J'en profite pour glisser un mot pour Darja Dubravcic dont le travail – et les souches fluo – ont contribué à donner un peu plus de relief et de sens au **Chapitre 5**.

En remontant un peu dans le temps, j'aimerais témoigner de ma gratitude envers Hayakawa-sensei pour son accueil, sa disponibilité et le gyutan, lors de mes deux visites à Sendai, mais aussi lors de mes fréquentes incursions virtuelles sur ses ordinateurs après mon retour à Lyon. J'en profite pour remercier Furuhashi-san, les Tsujino, Sato-san, Kimura-san, Omiya-san et Martin Robert pour leur accueil et leur gentillesse durant mes séjours.

Merci aussi à Christophe Anjard sans qui le **Chapitre 3** serait bien famélique et dont l'assistance a fini par me faire aimer et m'approprier cette partie de ma thèse qui, je peux bien l'avouer à présent, ne me semblait pas très excitante il y a trois ans.

Nous en arrivons aux deux personnes sans qui ces lignes n'auraient même pas de raison d'être. Il s'agit bien sûr de vous, Charlotte et Jean-Paul. Évidemment, un thésard doit toujours se plaindre de ses chefs, mais je ferai une exception ici. En effet, ces quelques grommellements ne doivent pas masquer la part immense de ce qui vous est dû dans ce travail, directement ou indirectement. Après une période d'observation réciproque, vous avez su me suggérer des idées, me laissant une grande latitude pour développer celles qui me semblaient les plus prometteuses tout en insistant discrètement sur celles qui méritaient selon vous plus de considération ; mais vous m'avez aussi laissé tracer une voie qui soit la mienne et pas celle d'un scénario préécrit. Je tiens à vous remercier tous les deux ensemble car vous avez vraiment eu des apports différents, complémentaires même, mais tous deux essentiels. Parmi les petites choses qui s'ajoutent à ces grandes idées, je n'oublierai pas, Charlotte, comment tu as remplacé mon coude fracturé pour terminer une série de manips, ou encore Jean-Paul ces quelques ramens dégustés en ta compagnie près de Sanjo-machi.

Dans la liste des "encadrants non-officiels", je me garderai bien d'omettre François Detcheverry, qui a apporté son regard un peu extérieur, d'une rigueur parfois énervante mais ô combien utile. Merci aussi à Alexandre Solon pour sa disponibilité et son ouverture lors de notre collaboration, et dont l'implication a dépassé les "simples" simulations pour faire émerger des idées pertinentes sur les interactions entre cellules.

Cela me permet de faire une transition habile vers son ex-grand chef, Julien Tailleur, qui n'a pas pris part directement à ce travail, mais dont les conseils et avis ont certainement joué un rôle décisif dans son avancée. J'ai énormément apprécié la façon dont tu donnes de ton temps pour guider des étudiants qui ne sont pas forcément les tiens, que ce soit dans leurs choix de carrière ou de sake. Il est important que des gens comme ça existent, je tenais à le souligner. Merci aussi à Hugues Chaté qui a co-organisé avec lui la Kyoto Winter School, moment charnière de ma thèse, et pour nos discussions lors de nos deux rencontres nippones. Merci à Vincent Hakim dont les commentaires avisés m'ont grandement aidé à mieux appréhender mes expériences du **Chapitre 4** et leur contexte plus général. Merci enfin à nos compères Kyotoïtes pour ces bons moments

passés ensemble, entre autres bien sûr Félix mais aussi Agnese, JFR[1] et Chloré[2] Barré.

Les chercheurs sont des gens qui font de la recherche, mais ce sont avant tout des gens, et la bonne ambiance qui règne dans la bi-équipe Bioliquides aux Interfaces Physiques contribue énormément à l'entrain que l'on met dans le travail. Merci donc à mes compagnons de paillasse : Antoine, sa bienséance confraternelle, ses poulpes et ses hymnes russes ; Sham, dont le nom fait toujours trembler de frayeur mes cellules ; Flo, son optimisme et son calme légendaires ; Tess qui m'a tenu compagnie dans le rush des dernières manips ; Angélique qui a amené un peu d'ordre dans la salle (ou bien est-ce la fin de mes manips ?) ; Hélène, Sylvain et Thomas, qui font semblant de maniper bien que tout le monde sache qu'ils sont permanents ; et les myriades de stagiaires qui ont squatté MA hotte plutôt que de me faire du café ou des massages – Peiyuan qui m'a aidé à initier les manips de colonies, Manon #1, Manon du ker-lac, Joseph II d'Angleterre, Steven et Larry qui tentent de prolonger un peu cette thèse, Souhila, Maxime, Cédric et Layla.

Merci à mes co-bureau pour l'équilibre concentration / bonne humeur que nous avons maintenu avec succès durant ces nombreuses années : Vasilica, Clara et sa boîte aux lettres, Marie-Julie l'intermittente, Antoine et Angélique voire Flo à un moment (cf ci-dessus), Séb et son humour de très bon goût, Marie-Émeline et ses questions de maths impossibles, Teresa à qui incombe désormais la lourde charge de s'occuper de la Balle, (Alexis, Catherine, Cora, Flo et Tess qui ont un temps partagé notre bureau,) et bien sûr Menka, à qui je ne pourrais témoigner de toute ma gratitude pour ces quelques années de voisinage sans outrepasser joyeusement le nombre de pages habituellement admis pour une thèse de physique.

Merci aux vieux qui ont posé les bases d'une ambiance que nous avons pu porter au firmament : Henri, Baud' dont les tacles ont épargné mes genoux, Isaac dont les pizzas ne m'ont, elles, pas épargné, Antoine encore, Pauline qui a beaucoup d'humour, Nico qui a aussi beaucoup d'humour "mais pas avec tout le monde" en plus de certaines aptitudes à la basse, et hop je case – ici, désolé si c'est un peu réducteur – un merci à Val' et ses biscuits à la maïzena, Lulu pour les belotes, les films thaï(wan)landais et les discussions improbables. Merci aux nouveaux qui ne sont plus si nouveaux, et qui ont assuré une relève de qualité, à commencer par ceux qui voient en ce moment fondre sur eux la deadline de rendu de leur "manuscrit" : Flo et Séb encore, Félix, compagnon de

[1] Jean-François Rupprecht pour l'état civil
[2] le prénom N'a PAS été modifié

foot et de yakitori qui a su prolonger la tradition pizzaïolesque, Loren la déménageuse narcoleptique et Alex dont on attend la démo d'aikido en apesanteur ; à ceux qui sont plus ou moins vraiment nouveaux : Ronan le footeux converti à la bière et au basket (enfin, surtout au basket), Quentin et son obsession pour le raifort, Manu, toujours présent pour alléger les pichets, Teresa et Marie-Émeline. N'oublions pas les post-docs et les américains qui ont illuminé Brillouin de leur passage éclair : Baptiste et ses grandes idées, Nico Bruot dont la source d'anecdotes ne tarit jamais, Pascal le grand frère, qui a initié la diaspora ILMienne vers le Chili, Rodney et ses rendez-vous musicaux, et Cora au langage ~~fleuri~~ niçois ; Calina pour la côte ouest, John Biddle et ses barbecues hivernaux pour la côte est.

Merci à ceux du Kastler aussi, et notamment : à Julien, même s'il a sûrement plus couru pendant les matchs du PSG chez lui ou au Bryan's que quand il était sur le terrain ; à mes nouveaux meilleurs amis, Antoine le brasseur et Nora la secrétaire ; à Daniel et son aura de gentillesse.

Cela me fait penser à remercier ceux avec qui j'ai réalisé le plus grand accomplissement de ma thèse : il s'agit évidemment de l'équipe victorieuse du Tournoi de Futsal des Personnels 2015 : Baptiste, Félix, Julien, Nico Terrier, Ronan.

J'ai bien sûr gardé le meilleur pour la fin : un grand KERCI aux ker-doctorants que furent Alexis, Chatherine, Flo, Menka et Simon pour cette studieuse ker-semaine ogienne, mais aussi pour le ker-lac, la ker-belette... Alexis que j'ai découvert sur le tard mais avec qui on a eu le temps de réaliser la ker-réunification des ker-bureaux, de développer moults ker-avions, de faire quelques ker-parties épuisantes de ker-ridor ball... Chatherine qui assortit toujours ses ker-idées les plus folles de sa ker-bonne humeur. Flo et nos grandes ker-discussions que nous tiendrons désormais en japonais. Menka et Simon et ce ker-voyage qui s'annonce... En passant, si l'un de vous lit ces ker-remerciements, où en est notre Phys. Rev. Postcard ?

Merci enfin à tous les permanents avec qui j'ai plus ou moins – jamais assez – interagi: Christophe et Laurent les jazzeux, Osvanny le basketteur, Rémy le barista, qui m'a bien aidé pour la litho, Agnès, Gilles, Christophe, Cécile, Jean, Marie et sa pancakière, Catherine, Anne-Laure et Mathieu ; à François et Jean-Michel pour leur aide sur le bain thermostaté ; aux informaticiens, Ruben, Fabio, Thomas et Carlos, pour leur assistance discrète mais précieuse ; à Christelle, Delphine, Sandra, Dominique et Audrey qui m'ont supporté dans les tâches administratives.

Cette thèse ne serait pas si je n'avais au préalable développé ma passion

pour la biophysique. Un grand merci à ceux qui m'ont accompagné dans cette démarche, que ce soit à l'X, au MBI, au Master PSB ou au LKB, et notamment aux "MEB non-officiels" : Œuphélie, Orly, Rémi et Yara avec qui nous avons passé une fabuleuse année de M2 terminée en beauté sur les plages basques.

Cette thèse, en outre, est indissociable de mon séjour à Lyon. Merci donc à Clo que je tiens pour responsable des prémices de mon amour (certes immensément inférieur à celui que je porte à Paris) pour cette ville. Merci à Mélo, puis Caca et Clém, et Claudio quelques mois, qui ont aussi contribué à rendre la vie tellement agréable au Garillan.

Puisque l'on parle de colocs, merci à celle qui m'a suivi à Lyon, pour qui j'ai traversé une rivière et un fleuve, puis qui m'a précédé à Paris, celle grâce à qui tout l'ILM est persuadé que je fais de l'aviron, et dont la présence m'a rendu plus facile l'acclimatation à la vie provinciale !

Merci à tous ceux qui, bien qu'étrangers à la biophysique et à Lyon, ont continué à accepter de me voir et à rendre la vie chanmé : Nabot et sa Marion, Tantine Gwen et Maumau les petits suisses, Ju la Normande du Limousin, Sanka et Celia les matouxx haut-savoyards (et Augustin, bien qu'il ait volé une partie de la vedette à ma soutenance), Jérém, Guigui, qui a, et Valentina, Nico, Fleur. Trüc de ouf !

Enfin, pour terminer par le début, je finirai en remerciant ceux qui m'ont toujours soutenu, de ma naissance à l'organisation de mon pot de thèse : mes parents, mes sœurs et mes grands-parents.

Mais pourquoi tout le monde met une citation au début de sa thèse ?

Dr Simon Gravelle

Contents

Preamble

Collective effects in the motion of cells have aroused a huge interest in the last decade, bringing together cell biologists, soft matter experimentalists and statistical physics theoreticians. However, a detailed account of the mechanisms, by which the interplay of single cell motility and cell-cell interactions can lead to collective behaviours – that can be very different from the individual ones – is still missing in most cases. The ambition of this thesis is to contribute in bridging the gap between the refined understanding of processes at the molecular scale allowed by the approach of cell biology on one hand, and the population scale observations of *e.g.* oncologists or developmental biologists on the other hand. This is why our focus is on the quantification of single cell behaviours in various collective contexts, with the constant aim of relating them to the macroscopic dynamics of the system.

The **Chapter 1** is an introduction to the concepts that are essential for the understanding of the following experimental and modelling work. It presents the basics of cell biology with a focus on the motility and cell-cell interactions, draws a comprehensive picture of the state of the art on collective effects in the cell motion, and introduces the rising field of active matter, whose framework we use to transition to the collective scale.

The main object we use to describe the properties of a single cell is their trajectory. In **Chapter 2** we give an overview of the usual descriptions of cell trajectories as random walks. Then we present the methods we used to obtain and analyse such data. Last we use computer simulations of self-propelled particles to calibrate our analysis tools.

We came to the study of such collective motion effects following the previous work of L. Golé, who demonstrated the existence of a chemical communication system in *Dictyostelium discoideum* cells that acts to down-regulate the single cell motility at the scale of the population. This particular system, which relies on a **Quorum-Sensing Factor (QSF)** is the central subject of **Chapter 3**. There we intend to decipher the functioning of this regulatory system.

The discovery of this feedback from the population raised the question of the output and potential interest of such a phenomenon, not only at the scale of individual motion but also for the group as a whole. To get insight into this issue, we developed a colony spreading assay, which is the topic of **Chapter 4**. In order to have a good control on the experimental parameters and to get reproducible data that we could analyse in a rationalised way, we used micro-fabrication to design the initial cell colonies. This work evidenced the fundamental role played by contact interactions, which allowed the cells to invade more efficiently the free space when the colony was initially denser.

The study on the QSF led us to observe cells plated on a large surface at high density. **Chapter 5** describes an unexpected aggregation phenomenon that occurs spontaneously in this situation. We quantified this observation, made in the early stage of this doctoral work, but the actual lack of control that we had on the system drove us to put it aside in favour of the more mastered experiments that precede in this manuscript.

In **Chapter 6** we draw the conclusions of this work and put them into perspective in the broader context of generic collective regulation of cell motility.

The appendices provide the reader with clues to go further into this work. **Appendix A** is just supporting material for the analysis of trajectories with two relaxation times in their direction. **Appendix B** introduces methods that are currently in a development stage to better characterise the aforementioned trajectories. Last, **Appendix C** is an in-progress extension of the work of **Chapter 4**, aiming at integrating the effect of **Contact Enhancement of Locomotion** that we uncovered into a mean-field description of the cell colony.

Résumé substantiel

Malgré une extraordinaire diversité, apparente et bien réelle dans la plupart des cas, l'ensemble du vivant repose sur des bases communes extrêmement conservées, de la plus petite bactérie au mammifère le plus grand ou le plus complexe. L'une de ces bases est la cellule : c'est l'unité élémentaire du Vivant, de même que l'atome peut être considéré comme l'unité élémentaire de la matière en général. Comme les atomes, les différentes cellules ont des constituants de base de natures similaires mais en nombre ou fonctions particulières variables ; comme les atomes, les cellules peuvent s'assembler de façon plus ou moins complexe pour former des amas homogènes ou hétérogènes : il s'agit des organismes pluricellulaires.

L'analogie, cependant s'arrête là. Les constituants de la cellules sont beaucoup plus complexes et beaucoup plus variés que les particules qui forment les atomes : c'est un ensemble très organisé de machineries moléculaires en solution aqueuse, isolé du milieu extérieur par une barrière physique. Cette dernière est le plus souvent une membrane constituée d'une double couche de lipides, parfois complétée par un mur cellulaire dans le cas des plantes, ou d'une seconde paroi de nature variable. Les machineries comprises dans cette barrière peuvent être classées en trois types principaux : la machine génétique, dont le composant de base est l'ADN qui contient l'information, assure l'hérédité des propriétés cellulaires au cours de la reproduction – mitose ou meïose – et constitue en quelque sorte le livret d'instructions de base pour le fonctionnement de la cellule ; les voies de signalisation sont les voies de communication à l'intérieur même de la cellule, et permettent à la fois la transcription et la traduction du code génétique en messages et en actions moléculaires, et l'intégration de signaux de toutes origines qui peuvent agir à tous les niveaux du fonctionnement cellulaire ; enfin, les effecteurs, la plupart du temps des protéines ou des complexes protéiques, réalisent des tâches variées allant de simples modifications chimiques à la génération d'actions mécaniques par exemple. Cet ensemble de machineries fait que la cellule est souvent qualifiée d'"usine" moléculaire, exécutant des tâches nombreuses, variées et complexes.

Une de ces tâches est la motilité, qui désigne l'aptitude des cellules à se mouvoir. C'est une propriété fondamentale de la cellule, qui intervient chez tous les organismes vivants, et à toutes les étapes de la vie. Certaines bactéries possèdent ainsi la capacité de nager grâce aux mouvements coordonnés de filaments, appelés flagelles, mus par des moteurs moléculaires. C'est le même type de système qui permet aux spermatozoïdes de se mouvoir jusqu'à l'ovule dans la reproduction sexuée des mammifères. Certaines cellules possèdent aussi la capacité de se mouvoir sur des surfaces, en se tractant *via* des déformations appelées protrusions, ou pseudopodes dans le cas des cellules amibiennes, objets de cette thèse. Une telle motilité est à l'origine des mouvements morphogénétiques menant à l'établissement de la forme des organes chez l'embryon, mais aussi impliquée dans la réponse immunitaire, la fermeture de plaies ou la dissémination de cellules cancéreuses à partir de la tumeur principale.

Ces phénomènes impliquent souvent plusieurs cellules en mouvement, voire en mouvement coordonné. Il est clair que la motilité de cellules dans un contexte collectif peut être affectée par la présence d'autres cellules, *via* des interactions de différentes natures. L'objet de cette thèse est de mieux comprendre la nature et les effets de ces interactions. Pour cela, nous avons adopté une démarche décrite brièvement dans les lignes qui suivent.

Tout d'abord, ce travail porte sur l'étude d'un organisme modèle : l'amibe *Dictyostelium discoideum*. Cet organisme, qui appartient à la famille des amibes sociales, se trouve à l'état naturel dans les sols forestiers. En présence de bactéries dont elles se nourrissent, les cellules restent à l'état dit *végétatif*. Elles restent unicellulaires, exprimant notamment peu de protéines d'adhésion cellule-cellule, se divisent, et se meuvent de façon aléatoire. En cas de carence de nutriments, elles rentrent dans un cycle de développement où elles commencent par s'agréger sous l'influence de chimio-attractants, puis se différencient en un organisme multicellulaire qui prendra plusieurs forment avant de culminer en un corps fructifère, sorte de champignon dont une partie des cellules forme des spores résistants. Ces derniers, une fois disséminés par le vent ou d'autre êtres vivants, pourront former de nouvelles colonies de cellules végétatives si les conditions s'y prêtent.

Dans le présent ouvrage, nous ne considérons que des cellules végétatives, capables de substituer un milieu de culture riche – le HL5 dans notre cas – à leur source naturelle de nutriments, les bactéries. Ces cellules ont la propriété de se mouvoir relativement vite sur des surfaces – avec une vitesse de l'ordre de quelques micromètres par minutes – et de façon

aléatoire. Leur mode de migration est caractéristique des cellules dites amibiennes, que l'on retrouve par exemple chez les leucocytes ou des cellules métastatiques d'un stade avancé. En outre, de nombreux sysèmes – jusqu'à génération des forces par le cytosquelette d'actomyosine par exemple – partagent des propriétés communes avec d'autres modes de motilité. Nous basons nos observations sur les trajectoires de cellules individuelles, dont nous décrivons les propriétés dans le cadre de modèles physiques.

Dans le **Chapitre Premier**, les différentes notions utiles à la compréhension du présent manuscrit sont introduites : bases de la biologie et de la motilité cellulaire, régulations connues de phénomènes comme la prolifération cellulaire à l'échelle d'une population, observations de mouvements cellulaires collectifs. En particulier, nous décrivons les principes fondamentaux de la matière active. Cette discipline récente s'attache à comprendre la dynamique, fortement hors d'équilibre, de systèmes constitués d'individus qui consomment localement de l'énergie tirée de l'environnement et interagissent éventuellement. Les particules autopropulsées, classe majeure de particules actives, sont ainsi des particules dont la propriété centrale est de se mouvoir selon une vitesse qui leur est propre, et en aucun cas dictée par des lois de conservation du moment ou de l'énergie. Elles constituent aussi, comme décrit plus loin, un potentiel excellent modèle pour des cellules motiles.

Dans le **Chapitre Second**, nous revenons en détails sur les propriétés des trajectoires cellulaires, ainsi que les moyens de les obtenir et de les analyser. Les différents types de marches aléatoires sont abordées, et notamment la plus simple marche aléatoire persistante, modèle de base pour les trajectoires cellulaires. Le lien qui existe entre les propriétés de ces dernières et des signaux provenant de l'environnement ou d'autres cellules sont aussi traités. Enfin, différentes manières d'analyser les trajectoires sont présentées, avec une attention particulère portée sur l'influence de l'aspect discret des données et de leur résolution, à la lumière de nos méthodes d'imagerie et de suivi cellulaire. Cette analyse se base notamment sur le traitement de trajectoires simulées numériquement, et dont les propriétés s'approchent de celles que l'on attend pour des cellules réelles.

Dans le **Chapitre Tiers**, nous étudions un système de régulation de la motilité *via* un facteur chimique sécrété, ou facteur de quorum (QSF en anglais). Nous avons ainsi caractérisé sa nature chimique, la dynamique de sa sécrétion et de la réponse cellulaire, ainsi que les voies potentielles de signalisation qu'il induit, et ses effets au niveau morphologique et fonctionnel sur les cellules individuelles. Notamment, nous démontrons

qu'il s'agit d'une macromolécule neutre et de nature non protéique qui induit une réponse *via* des récepteurs membranaires et des protéines G. Nous montrons aussi que la dynamique de sécrétion est auto-régulée et pourrait constituer un système de mesure précis et robuste de la densité cellulaire.

Dans le **Chapitre Quart**, nous plaçons les cellules dans une situation stéréotypée d'étalement de colonie. Plus précisément, à l'aide de techniques de micro-fabrication, nous contraignons un nombre contrôlé de cellules dans un disque de diamètre fixe avant de relâcher la contrainte. Cela constitue un modèle basique et reproductible de colonie cellulaire, qui pourrait mimer par exemple de façon très simplifiée l'échappement tumoral. Ce système nous permet de mesurer à la fois des grandeurs macroscopiques comme le rayon de la colonie où la densité cellulaire, mais aussi microscopique à travers l'étude, une nouvelle fois, des trajectoires individuelles. Cela nous permet de distinguer plusieurs régimes d'étalement. Aux temps longs (de 10 à 48 h), la prolifération cellulaire domine la dynamique, comme prévu par les modèles classiques. L'effet du QSF caractérisé au chapitre précédent est aussi visible à cette échelle. Aux temps plus courts, au contraire, nous avons mis en évidence un effet jusqu'à présent non rapporté des interactions de contact. Ainsi, l'étalement est d'autant plus rapide que la densité cellulaire est élevée. Cela est lié à l'apparition d'une polarisation des cellules vers l'extérieur de la colonie, sous l'influence des interactions qui accroissent la persistance du mouvement. Nous formulons à la fois un modèle de particules uniques et des équations de champ moyen qui rendent compte de ce phénomène, et nous discutons les mécanismes possibles régissant ces interactions.

Dans le **Chapitre Quint**, nous rendons compte de l'observation d'un comportement collectif observé lorsque les cellules sont déposées à haute densité. Elles forment alors des agrégats, probablement tridimensionnels, mobiles et dynamiques. Nous décrivons leur dynamique de formation et leur distribution en taille, avant de formuler des hypothèses quant aux principes physiques et biologiques menant à leur apparition.

Dans le **Chapitre Sixte**, enfin, nous revenons sur les résultats principaux décrits dans le présent ouvrage, en détaillant les conséquences et les perspectives qu'ils ont ouverts.

1. Introduction

Cells move. The consequences of this very simple assertion are huge: without cell motion, the early shaping of multicellular embryos should be completely reprogrammed. Supposing that tiny difficulty to be overcome, basically no physiological function could be carried out. Any infection? Let the infected tissue struggle with it: the body's policemen, also called leukocytes, would be stuck at home. It might possibly be an issue, as immobile epithelial cells might not be able to close any wound. Anyway, you would not realise it, as your neurons would not be able to make any plastic connexion that would endow you with any cognitive capacity. Besides these and other interesting abilities that cell motion confers to living organisms, it is also involved in the development of pathologies. Among them, the most commonly cited is certainly cancer, which involves a great deal of cell motility in the invasion of surrounding tissues and the formation of metastases.

For these reasons, biologists have studied cell migration for decades, and produced a lot of knowledge about its molecular functioning at the single cell level. As far as physicists have been interested in biology, they have also contributed to this by bringing their singular view on forces and motions. Yet, none of the relevant phenomena that rely on cell migration involves isolated cells. Indeed, *in vivo*, cells always encounter other cells in their vicinity, hence cell-cell interactions play an important role in these processes. Certainly, collective behaviours of moving agents have been gaining a great interest for 15 to 20 years, at the same time in the communities of cell biology, physics, and at their interface. However, the general mechanisms by which single cell motility is affected by cell-cell interactions to give rise to group effects remain elusive. In this chapter, we introduce the foundational concepts that are necessary to our biophysical approach of this difficult question, which we develop further in the present manuscript.

Contents

1 The cell, atom of the living matter

Despite their huge diversity in size, shape, complexity or environment, all living organisms are built on the same basic unit: the cell. Despite, once again, some differences, all the cells have extraordinarily shared properties. For instance, the great majority of cells are typically from 1 to a few tenth of microns (µm) in size, whether the whole organism is invisible to the bare eye or as big as a whale. The structure and molecular basis of the cell are also among these common features: they are small bags of water – or rather of complex aqueous mixture – separated from the outer space by at least a plasma membrane made of a lipid bilayer. A cell can be an individual organism by itself: this is the case for all the bacteria and archea, but also for most of the eukaryotes; cells can also assemble into various multicellular structures to achieve functions that are not accessible to isolated cells.

Of course, the cells themselves perform multiple tasks. The information needed for their functioning is encoded in DNA, a sequence of nucleotides that is replicated when the cells divide, ensuring the continued existence of species. The DNA is transcribed in messenger RNA, which are subsequently translated in proteins using the genetic code, another strictly conserved property of all the living kingdoms. The proteins are the main effectors of various processes thanks to their specific and powerful catalytic properties, but the actual cellular processes are much more complicated than this simple linear sketch: in fact, a cell is similar to a complex microscopic machinery, with a well-defined spatial organisation and communication pathways that serve the regulation of all those processes (Fig. 1.1).

All these functions are carried out at the molecular level. This is the usual approach of cell biology to tackle them at the same scale: in this scope, the – oversimplified – strategy to study a particular phenomenon consists on finding mutants that behave unusually, regarding this phenomenon; then identifying the proteins they do not produce correctly and looking for their interaction partners, their cellular location... In this manuscript, we adopted the physicist's converse view: we consider the cell as the elementary unit, hence integrating all the subcellular events in their observable cell scale outputs. Then, the question is whether it is possible to interpret the collective dynamics as the product of the individual behaviours and the mutual interactions of the single cells. This way, the cell can be viewed as the 'atom' of living matter.

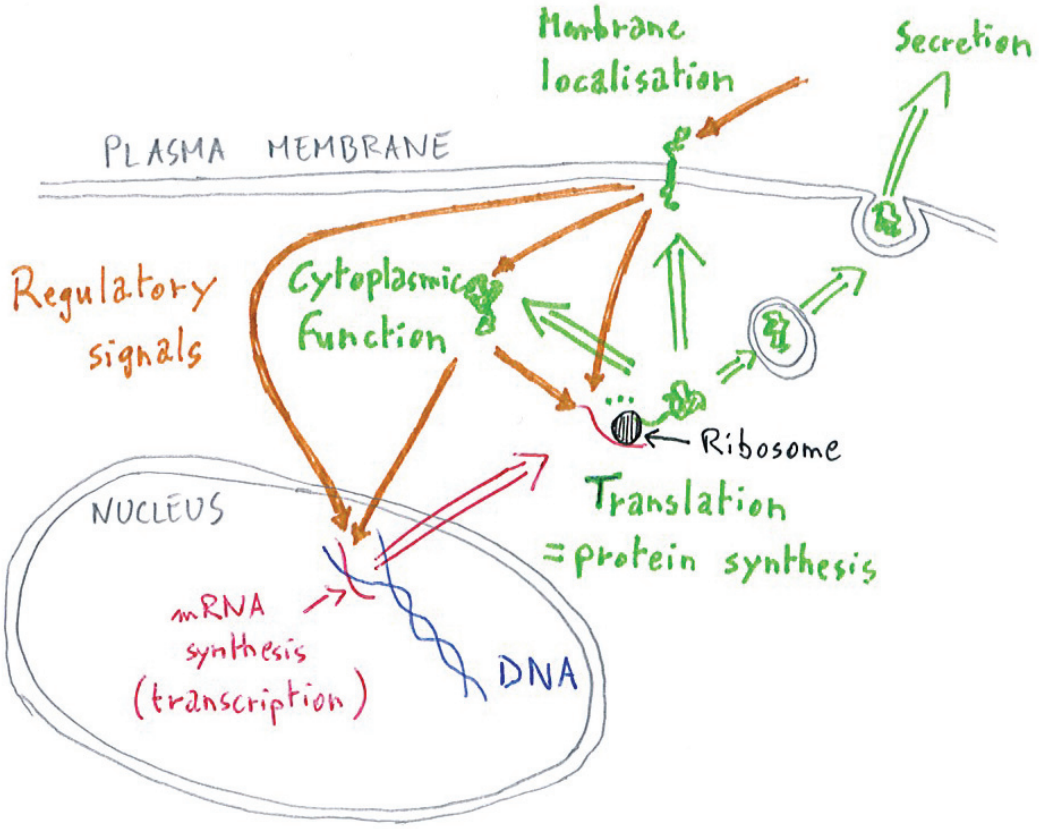


Figure 1.1: Dynamical architecture of the cell.
This cartoon presents an simplified view of a cell: DNA in the nucleus is transcribed in messenger RNA, which in turn is translated in a protein. The protein then carries out its function, either in the cell, within the external medium or at the border between the two. All the processes are regulated by at network of signalling pathways.

2 The cellular "vivre-ensemble"

Just as atoms can gather into molecules and complexes with a broad variety of properties, in certain conditions cells build multicellular structures. The latter can take the form of simple clusters of similar cells: for instance this is the case of laboratory spheroids [1], which are just aggregates of adherent cells that spontaneously become spherical under the action of effective surface tension. Here, all the cells play the same role, and there is no spatial structuring. Such spatial structuring occurs in *e.g.* biofilms or *Dictyostelium discoideum*'s development (Fig. 1.2). In the latter case, the cells, which are originally similar individuals, specialise in two main cell types: about 80% of the cells become spore cells and will be able to eventually grow new colonies, while the remaining

ones become stalk cells that only provide a structure for the others.

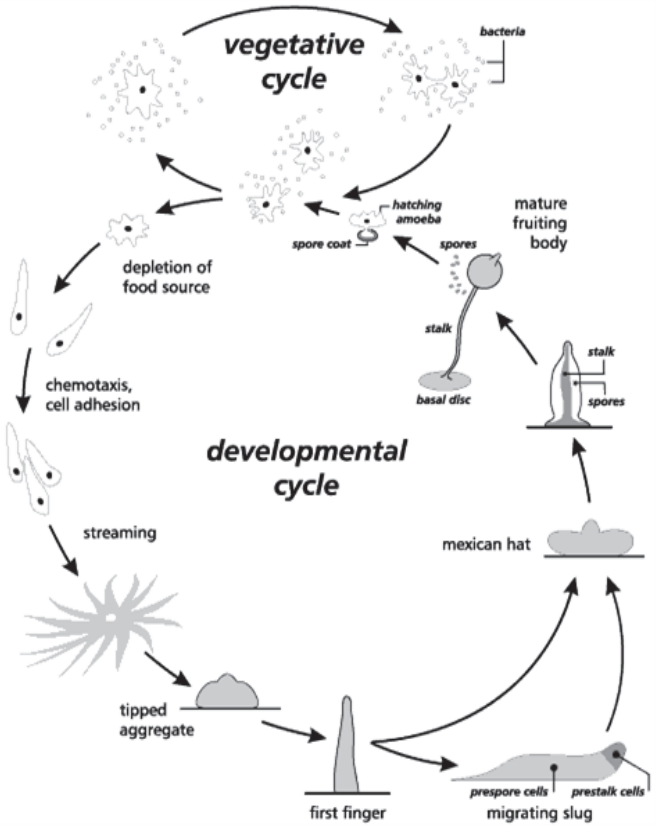


Figure 1.2: Dictyostelium discoideum*'s life cycle. In presence of food, the cells grow and move randomly, as unicellular, non-adhesive individuals (vegetative cycle). Upon nutrient starvation, they enter a developmental cycle: first the cells aggregate thanks to self-chemoattraction, then they differentiate in two main cell types: prespore and prestalk cells. The 'multicellular' body undergoes multiple morphological changes before the culmination of a fruiting body, from which the spores can be transported to a new niche, and where the stalk cells will eventually die.*

This is a simplified version of what happens during the development of truly pluricellular organisms: pluripotent cells differentiate in cells with specific capacities, which organise spatially according to defined patterns. This eventually gives rise to a potentially very complex organism composed of various organs that carry out specific tasks. Each of these organs is composed of various tissues, which are made up of cell populations and extracellular medium. To take the example of the skin, it is composed of three main layers: the epidermis at the outside, then the dermis and the hypodermis inside; the epidermis is itself made of a tri-dimensional epithelium of keratinocytes, with dispersed melanocytes, Langherans cells and Merkel cells (Fig. 1.3).

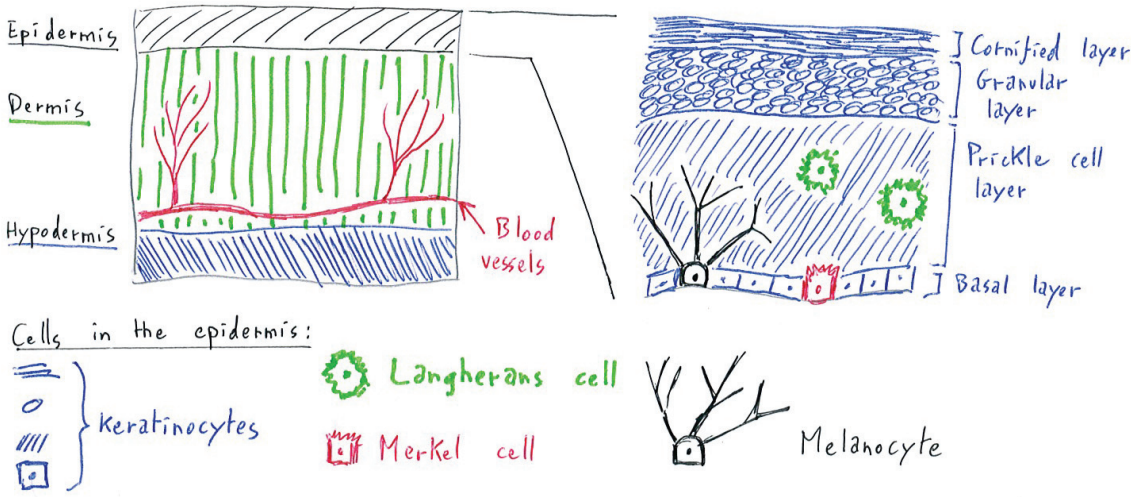


Figure 1.3: Organs are just structured cell populations in interaction: the example of the skin.
The skin is made up of three main tissues (left): epidermis in contact with the outer space, then dermis and hypodermis. Epidermis is itself a collection of interacting cell populations of various types(right): connective layers of epithelial keratinocytes, in which populations of melanocytes, Langherans cells and Merkel cells are sparsely embedded.

2.1 The two senses of the cell

In this same view, a body can be viewed as an assembly of interacting cell populations. Moreover, there are also interactions inside these populations, between cells of either the same or different types. These cell-cell interactions are fundamental in the functioning of the organism. To conform to easy anthropomorphism, think of a group of people without any perceptive ability: it is hard to imagine them act in coordination. Actually, it would be impossible for such an individual to react in any way to its environment. Similarly, the cells need systems that enable them to sense their environment. This happens at the molecular scale, most of the time through transmembrane proteins (Fig. 1.4). These proteins have three domains: a central hydrophobic one that anchors them to the lipid membrane; an extracellular domain that is able to sense external signal, for instance by binding to specific chemicals; a cytoplasmic domain, that interacts with other cytoplasmic molecules, hence transmitting the signal into the cell. In some cases, the integration of external signals may also be made through channels, that allow specific molecules into the cell, or by internalisation of external medium in endosomes.

In the following we focus on receptor-mediated sensing, to review the

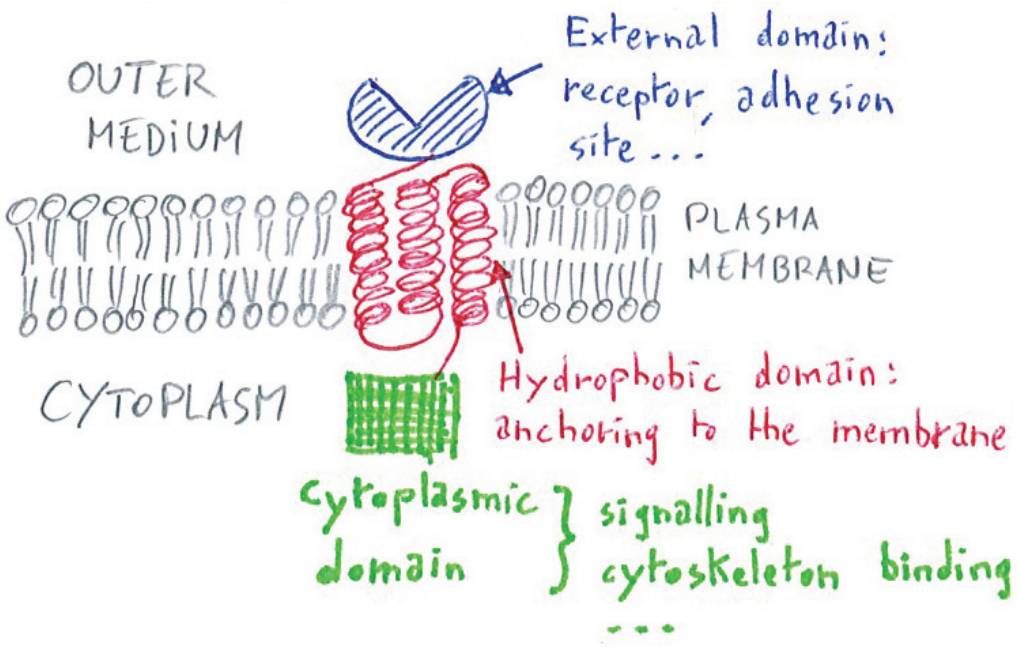


Figure 1.4: Membrane protein.
A membrane protein is a protein that is anchored to the membrane by a hydrophobic part (here, 3 red a-helices). Transmembrane proteins, which have both an external domain (blue) and a cytoplasmic one (green), are able to transduce signals through the cell membrane.

concepts of cell-cell interactions (Fig. 1.5). It is easy to imagine that membrane proteins provide the cells with a sense of touch: in that role for instance the cadherins, a family of cell-cell adhesion proteins, are very important. The neural crest cells use N-cadherin to detect contacts with other cells, inducing a mechanism known as Contact Inhibition of Locomotion (CIL, see section 4). In epithelial tissues, the adherens junctions are used to communicate mechanical signals: cells pull or push each other, and the forces are transmitted through E-cadherin to the cytoskeleton, inducing various responses. Yet, this contact sensing would not be sufficient: to many extents, sensing the others, further than one's closest neighbours, is of great importance for individuals in a population. To come back to our anthropocentered analogy, think of a group of both blind and deaf people. Although some information can be transmitted along the group by a sequence of physical contacts, as we will see in Chapter 4, it is hard to imagine any efficient coordination of the whole group in this delicate situation. To communicate with distant individuals, the cells are able to send signals away by secreting messenger molecules. These molecules can be transported by diffusion, or in some cases by convection in a fluidic system such as the blood or lymphatic vessels, and then detected by receptors on distant cells. Depending on

the systems, the information conveyed can be more or less complex. As we will see in Chapter 3, the overall cell density or cell number in a population can be encoded in the concentration of the signal. The cells may even integrate directions by sensing gradients, and this is when their active and dynamic nature becomes fundamental: in a gradient of external signal, the receptor can distribute over the cell body in a way that reflects this gradient, hence inducing a polarised intracellular response [2]. In the particular case of chemotaxis – the ability to move up or down chemical gradients – the directionality can also be detected even if the length scale of the gradient is larger than the cell body, by integrating signals in time and biasing the motion accordingly [3].

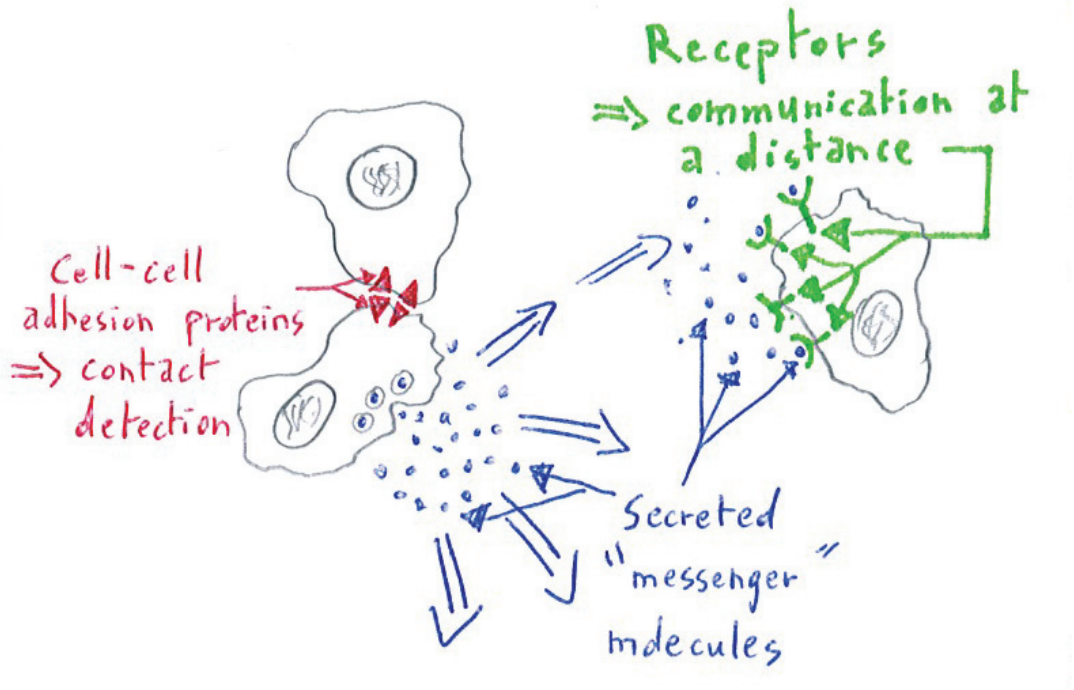


Figure 1.5: Cell-cell interactions.
Cells can sense specifically their neighbours thanks to adhesion proteins (red, left). They can also communicate at a distance through secreted molecules that are detected by the suitable receptors.

To sum up, the cells possess the ability to sense their neighbours in various ways, either at contact or at a distance, involving the 'knowledge' of various cues: directions, cell density or number, forces, induction of any activity (Table 1.1)... The response to these signals can be of multiple nature. Here we will focus on their effects on the two main processes that shape cell populations: growth and motility.

Sense \ Info	Contact	Distant	Number	Precise location	Direction
Hearing	–	✓	∼	–	✓
Sight	–	✓	✓	✓	✓
Smell	–	✓	∼	–	–
Taste	✓	–	∼	∼	N/A
Touch	✓	–	–	✓	✓
Contact sensing	✓	–	–	✓	✓
Quorum sensing	–	✓	✓	–	–
Chemotaxis	–	✓	∼	–	✓

Table 1.1: Comparative table of human and cellular senses. Property detected (✓), partly detected (∼) or not detected (–) by senses. In their ability to sense the spatial organisation of their neighbours, the cells would be similar to blind people (the taste is not really relevant to this capacity). They can sense neighbours precisely at contact, but at a distance they have only clues about their overall direction and number, and not their precise location.

2.2 Cell growth in a social context

As stated by Jacques Monod in his seminal article on bacterial growth [4], " it would be a foolish enterprise, and doomed to failure, to attempt reviewing briefly a subject which covers actually our whole discipline". Our discipline is not microbiology, but the ideas of this article, including this quotation, also hold in the case of eukaryotic cells. Trying not to be too foolish, we will try to give a brief overview of the underlying principles and molecular bases of cell proliferation, which in essence takes place in a collective context.

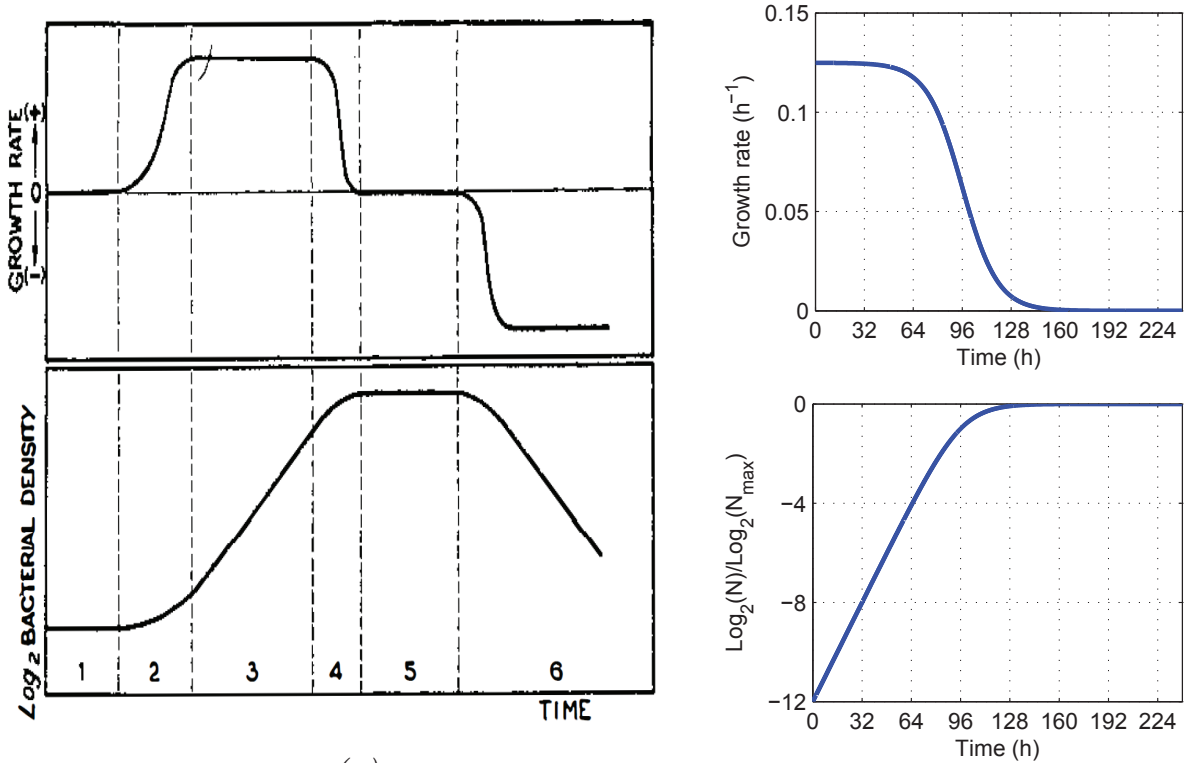
Rigorously, two kinds of growth must be distinguished: mass growth, and number growth. While the former refers to the increase in quantity of cell material, due for instance to duplication of DNA, production of proteins, or even simply water uptake [5], the latter represents the increase in *number* of individuals, due to cell division. The two can be different in some cases: when division but not duplication is inhibited for instance, it gives rise to the growth of individual, multinucleated cells, but not of the cell number. Here we consider the **number growth**. We also implicitly suppose the ideal case where cells divide correctly, keeping a constant mass per individual in average. Then, both growths are identical, and

in particular the density and packing fraction are equivalent.

Forgetting about various lag phases (Fig. 1.6a), there are two main, limit regimes in cell population growth: an exponential phase and a saturation phase. In the former, the cells divide at a constant rate, giving rise to an exponential growth of the cell number (1.1):

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \frac{n \ln 2}{t_2} \Rightarrow n(t) = n_0 2^{\frac{t}{t_2}} \tag{1.1}$$

where t_2 is the doubling time, or, $n(t) = n_0.e^{\lambda.t}$, using the growth rate $\lambda = \frac{\ln 2}{t_2}$ instead of the doubling time. In the latter the cells stop dividing at a given density. Of course the transition between the two regimes is smooth: it is called the "retardation phase", in which the division rate continuously decreases toward zero (Fig. 1.6).



(a) *(b)*

Figure 1.6: Typical cell proliferation dynamics.
(a) Sketches from [4]: (1) lag, (2) acceleration, (3) exponential, (4) retardation, (5) stationary and (6) decline phases. Growth rate (top) and corresponding evolution of the logarithmic density (bottom). (b) Computed growth and growth rate from expression 1.3 with $t_2 = 8$ h. It represents well the exponential, retardation and saturation phases (phases 3 to 5 in a).

This whole dynamics is quite well captured by the logistic growth model initially introduced by Pierre-François Verhulst in 1838 (see Ref. [6] and Equation 1.3). In this model the growth rate linearly decreases when the density approaches a saturation value called the carrying capacity K. The resulting time-evolution of the cell density is very similar to that of observed cell proliferation (Fig. 1.6): starting from a low density, an initial exponential growth phase is followed by a slowing down and a saturation

$$\frac{dn}{dt} = \lambda n(1 - \frac{n}{K}) \tag{1.2}$$

$$\Rightarrow n(t) = \frac{K}{1 + e^{\lambda(t-t_0)}} \tag{1.3}$$

Both phenomena originate from molecular processes that occur inside the cells. The exponential growth relies on the fact that the doubling time of the population, hence the cell division time, is constant. It is mostly due to the precise regulation of the cell cycle, as revealed in Refs [7, 8]. In axenic strains of *Dictyostelium discoideum*, the mitosis is short, lasting for about 10 minutes, and the cells spend most of the time in the S-phase (DNA replication phase) or the G2-phase (cell growth after DNA replication), giving rise to a narrow distribution of cell cycle times around 7h to 9h, depending on the strain.

The cell cycle is slowed down at high density until complete arrest. This slowing down implies that the individuals are able to sense the surrounding density. Apart from indirect effects as the medium acidification, or reaction to cellular waste or nutrient depletion, this is allowed by secreted communication molecules called "quorum-sensing factors". It is known for example that the secreted proteins AprA and CfaD act together to slow *Dictyostelium*'s proliferation by lengthening the G2 phase [9, 10, 11]. In their absence, the cells proliferate faster and reach a higher saturation density [9].

The regulation of the proliferation is a first example of the cellular "vivre-ensemble". In what follows, we will see that this notion is also very important for another cell property: the motility.

3 Cell motility

The term **motility** refers to the very important ability of cells to move actively. Using a machinery that we describe below, they can propel themselves at speeds ranging from a few microns per hour to a few microns per second for the fastest. This motion takes place either on a solid substrate – crawling motility – or in a liquid medium – swimming motility. The second is particularly relevant to bacteria and sperm cells; we will focus rather on the first type, widely observed in eukaryotic cells and especially inside multicellular bodies.

Indeed, cell migration is central to many phenomena, either normal or pathological, at various stages of one's life. During the development, large-scale cell movements, such as the gastrulation or the neural crest migration, shape the embryo and the future organs. All life long, in the intestine epithelium, the cells proliferate at the base of villi, then they differentiate during their migration toward the top where they eventually undergo apoptosis. The migration of single cells is important *e.g.* for immunity: the dendritic cells alternate between medium uptake and displacement phases to scan the potential foreign bodies present in the tissues [12, 13], and the inflammatory response involves leukocytes that move fast to chase the infectious cells. Also, the high motility of cancer cells is a real health issue, as this is what provides them the ability to invade the surrounding tissues and even to form new tumours, called metastases, far from the original one, which are one the main reasons that make cancer so deadly [14]. In the case of unicellular organisms, the motility is very important for ecological concerns, as it is a driving process in the dispersion of species in their environment.

3.1 The molecular bases of cell motility

Even though we will try to consider motility as an intrinsic property at the cell scale, it is important to lay its foundations, in order to be able for instance to interpret the output of interactions.

The crawling motility of eukaryotic cells mainly relies on the dynamics of the acto-myosin cytoskeleton (Fig. 1.7). The latter term refers to long linear polymer chains, assembled in various structures, that give the cell its architecture and mechanical resistance. Actin microfilaments constitute one of these structures, along with microtubules and intermediate filaments. It consists in a linear assembly of actin monomers, with a polarised activity: at one end, monomers are added, while at the other end they are removed by active processes under the drive of ATP (Fig. 1.7a).

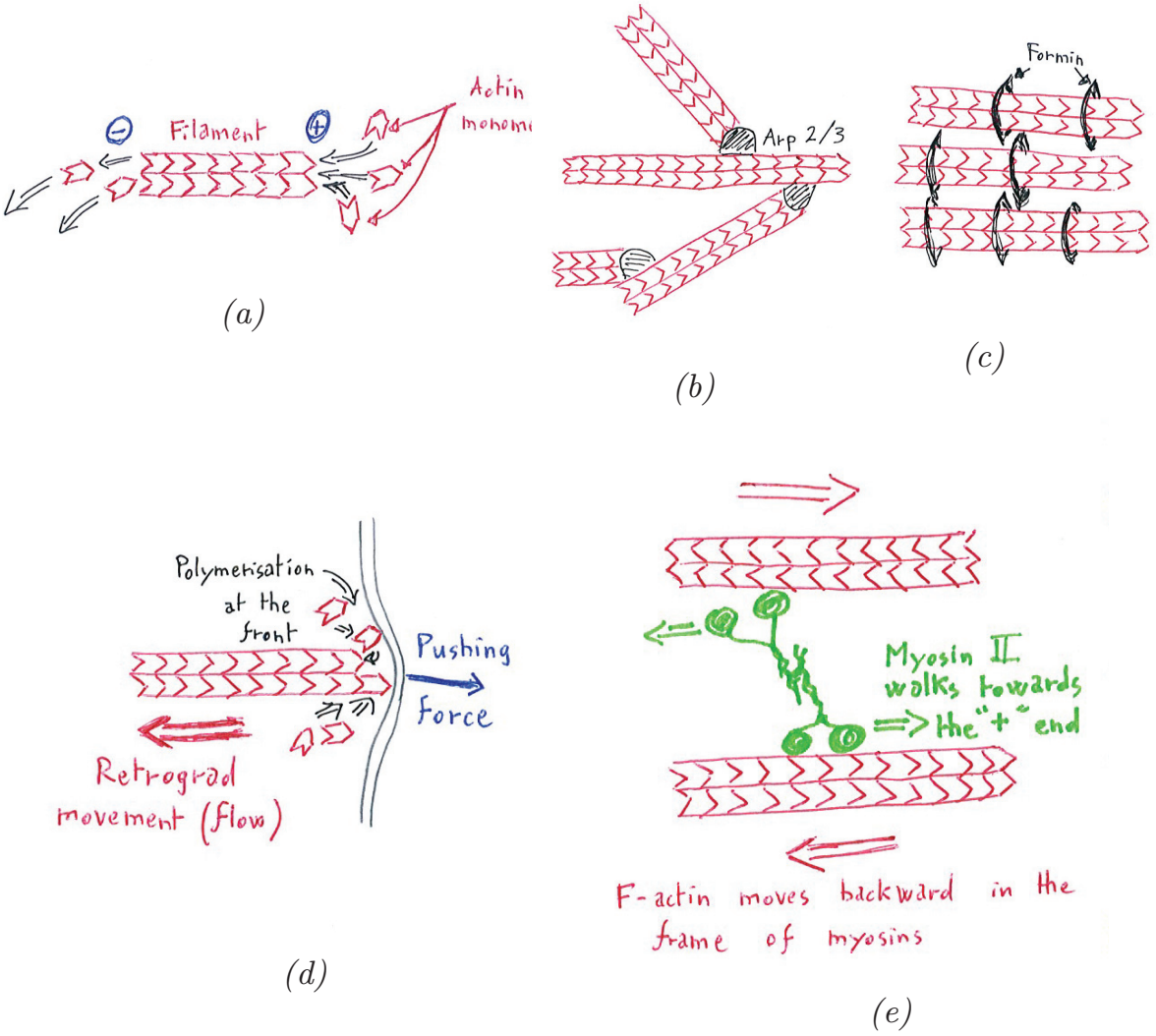


Figure 1.7: Actin structure and dynamics.
(a) Actin is polarised. Actin filaments polymerise at the barbed (+) end and disassemble at the pointed (-) end. (b,c) Actin can build supra-filament structures: branched actin (b) or parallel cables (c) for instance. (d,e) Actin can exert forces and motion: the polymerisation against the cell membrane induces a retrograde motion of the filament and a pushing force on the membrane if the filament encounters resistance. (d) anti-parallel bundles bound with myosin II motors have a contractile activity due to the relative motions of the filaments (e).

These dynamics give rise to forces and movements: if a filament polymerises, for instance, against the cell membrane, it will exert a pushing force on the membrane, and at the same time induce a retrograd flow of the continuously growing polymer (Fig. 1.7d). In addition, filaments can assemble in various structures thanks to cross-linking molecules, to form large bundles or flat extended meshes (Fig. 1.7b,c). Among these cross-linking molecules, molecular motors such as the myosins possess a special ability: they have two 'feet' that allow them to walk along the

actin filaments. These feet are linked to two α-helical coiled coil chains, which can bind either to a cargo they can transport, or to another myosin head. In this case, two linked myosins walking each on a different filament in opposite directions can give rise to a relative movement of both filaments (Fig. 1.7e). This provides the cell with a contractile ability, which is used for instance in muscle contraction, but also for single cell motility.

The actin cytoskeleton is connected to the outer space through transmembrane adhesion complexes. Thus, some cells can make focal adhesion to adhere to extracellular matrix molecules such as collagen or fibronectin; some others, among which *Dictyostelium discoideum*, stick to the substrate with non-specific adhesion. The turnover of these adhesion points and the intracellular actin remodelling may lead to a net motion through the following scheme (Fig. 1.8): the cell uses the adhesions to exert traction or friction forces on the substrate by polymerising actin bound to these adhesions; then the cell uses its contractile activity to retract its rear. Although simplified, this schematic view is quite general to all crawling eukaryotic cells and allows a good understanding of the properties of cell kinematics.

3.2 Single-cell motility as random motion

As we have just seen, cells are put in motion by a symmetry breaking between a protrusive front and a retractile rear. This implies that the cell is polarised, and that the dynamics of the motion relies on the dynamics of this polarity axis. Of course, external signals such as the chemoattractant that we evoked above can bias the polarity so that in average the cell moves towards the higher concentrations. However, even in the absence of any directional cue, protrusions still form and they last for a finite time. This defines the cell's **persistence time** τ_p, which is the time during which the cell moves with a memory of its previous direction. Of course, this does not last for ever in an isotropic medium and the cell eventually changes its direction of motion. This results, at a long enough time scale, in an apparent random motion: in particular, the asymptotic behaviour of the mean-squared displacement, or MSD, is linear in time:

$$\begin{aligned} \mathrm{MSD}(\delta t) &= \langle \delta x^2(\delta t) \rangle \\ &= \langle \|\mathbf{x}(t+\delta t) - \mathbf{x}(t)\|^2 \rangle \\ &= 4Dt \text{ for } t \gg \tau_p \text{ in 2D.} \end{aligned} \tag{1.4}$$

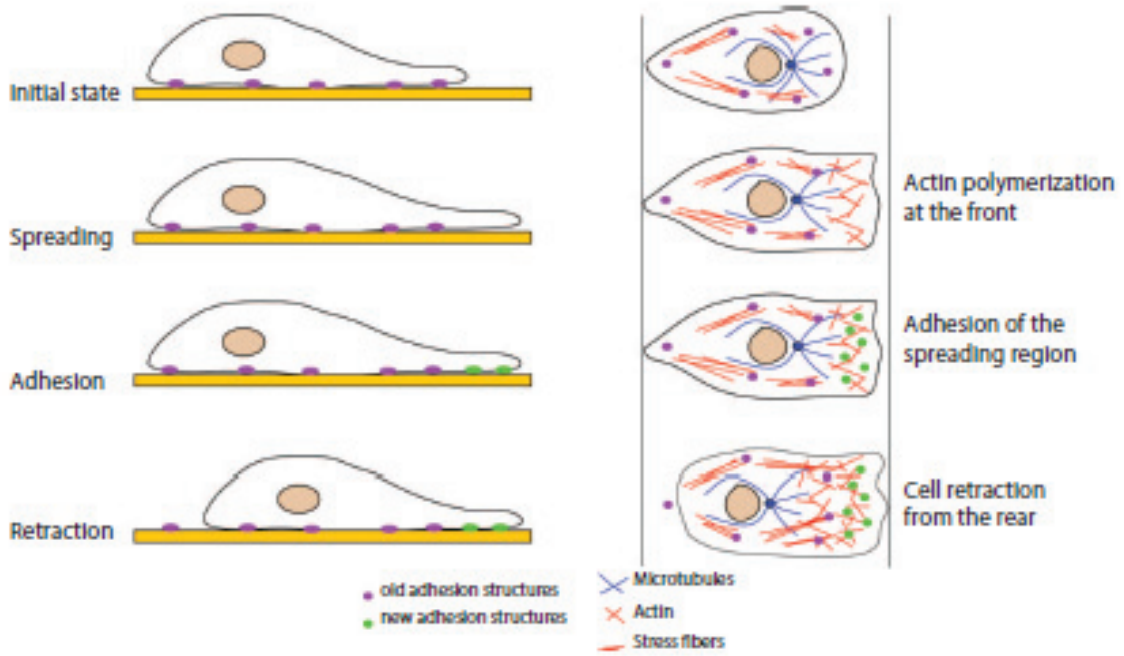
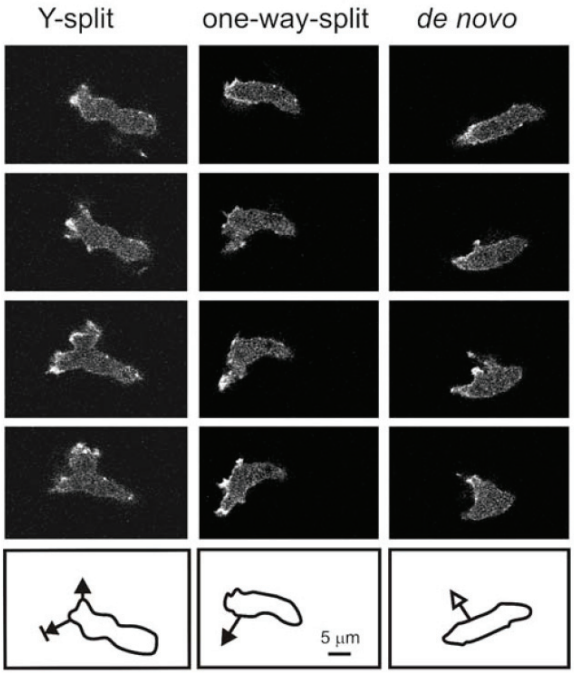
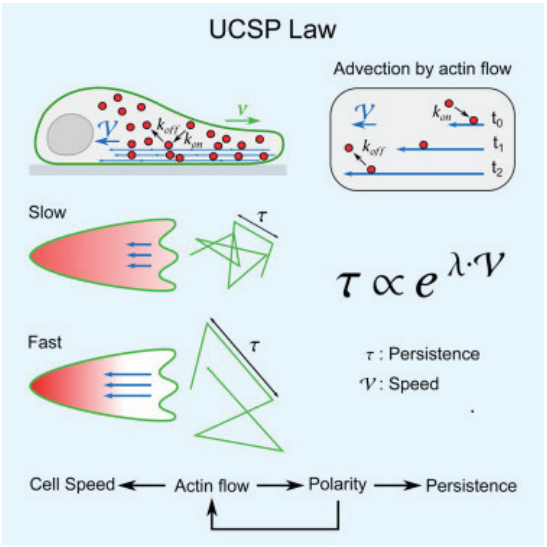


Figure 1.8: Cartoon of the migration mechanism of a cell on a 2D substrate (from L. Golé, PhD thesis).
First, the actin polymerisation pushes the membrane at the front. Then, adhesions are made so that this newly extended protrusion can exert forces on the substrate. Last, the rear retracts under the action of contractile actomyosin that detaches the adhesions in the back.

This allows to define an effective **diffusion constant** D, which characterises the cell motility at large time scales. In the $t \to 0$ limit, the motion is ballistic, at a speed v that is set by the protrusion-retraction activity. From that, it appears that the time scale at which the motion is observed and the trajectories sampled has a strong effect on the measurements that are made. This is the main topic of the **Chapter 2**. Also, understanding the dynamics of the cell polarity is fundamental to well describe the trajectories.



(a)



(b)

Figure 1.9: The cell motion is polarised.
(a)Dictyostelium cells follow their pseudopods (from [15]). The time-lapse pictures show the actin-filled pseudopods, that can originate from a previous one (Y-split, the bottom one then retracts, or one-way split) or be newly formed. In all cases, the cell turns in the direction of the new pseudopod. (b) Universal coupling between speed and persistence through actin-bound signals (from [16]).

Cell polarity. Bosgraaf *et al.* [15] observed that *Dictyostelium* cells produce pseudopods that grow for 11 s on average before shrinking when another pseudopod takes the lead (Fig. 1.9a). The latter can be either a 'splitting' pseudopod, that appears close to the pre-existing one and induces a persistent zigzag motion, or a *de novo* pseudopod, that forms at a random location along the cell body, leading to a more sudden reorientation of the cell. In this system, the ratio of *de novo* over splitting pseudopods sets up the persistence of the trajectory.

At a more intracellular scale, the polarity is known to correlate strongly with the mutual localisation of small GTPases of the Rho family: Rac1 and CdC42 are found rather at the front and promote the protrusion activity, while RhoA, which is related to the contractility, locates mainly at the rear. Recently, Maiuri *et al.* [16] proposed a physical model in which such polarisation cues are transported to the cell back by the actin retrograd flow (Fig. 1.9b). This model explains a coupling between the speed and the persistence of the cells, which has been observed in some cell types. It also predicts that the polarity axis can move in two main ways: first the direction can change smoothly, due to the fluctuations of orientation of the active protrusion; second, it can change abruptly by a sudden loss of polarity and building up of a new protrusion (Fig. 1.10). These two reorientation mechanisms, which we term respectively 'angular diffusion' and 'tumble', constitute the basis of active particle models, as we will see in **Section 5**, and they provide a quite complete set of tools to describe actual cell trajectories.

Until recently, the Ornstein-Uhlenbeck process [17] has been the simplest and most widely used model for cell trajectories. It describes a correlated random walk, for which the velocity-autocorrelation function C decays exponentially with time

$$\begin{aligned} C(\delta t) &= \langle \mathbf{v}(t+\delta t) \cdot \mathbf{v}(t) \rangle \\ &= \langle v^2 \rangle e^{-\frac{\delta t}{\tau_p}} \end{aligned} \tag{1.5}$$

and the MSD has an explicit expression called Fürth's formula:

$$\mathrm{MSD}(t) = 4D \left(t - \tau_p.(1 - e^{-t/\tau_p}) \right). \tag{1.6}$$

In **Chapter 2** we will review some more refined models, but the previous lines contain already the general concepts that are really essential to apprehend the properties of cell motion: the existence of a speed v at short times and a diffusion constant D at long times, with a switch of behaviour around the persistence time τ_p; and the corresponding short time ballistic regime in which MSD $\approx (vt)^2$, the long time diffusing regime with MSD $\approx 4Dt$, and the exponentially decreasing correlation function $C(\delta t) \propto \exp(-\frac{\delta t}{\tau_p})$.

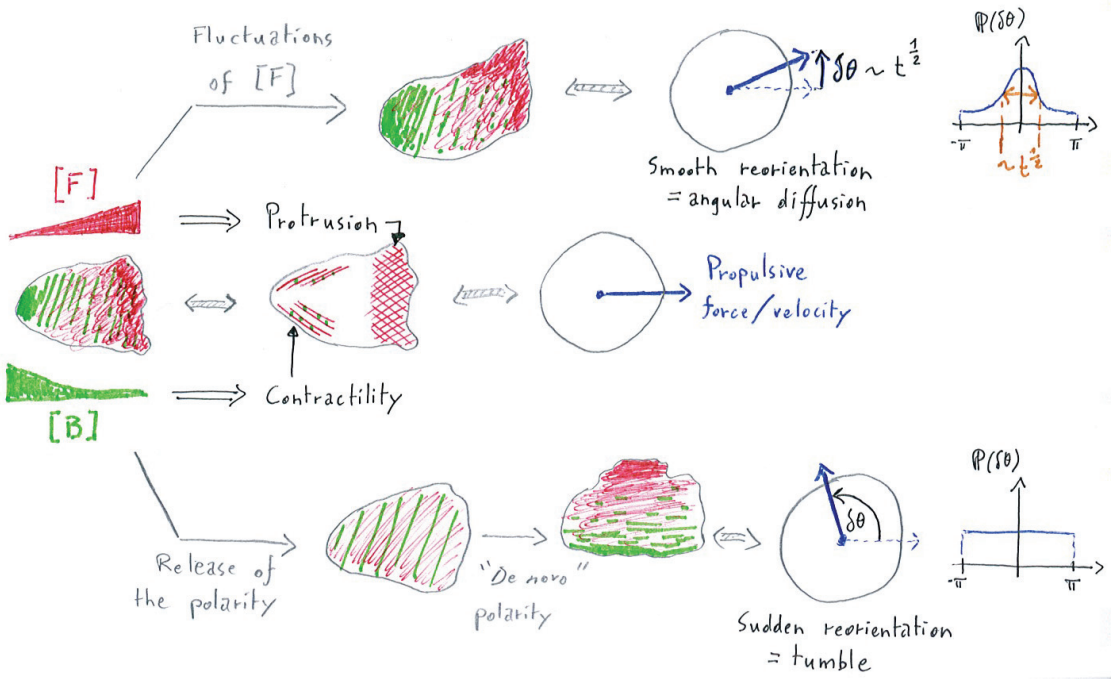
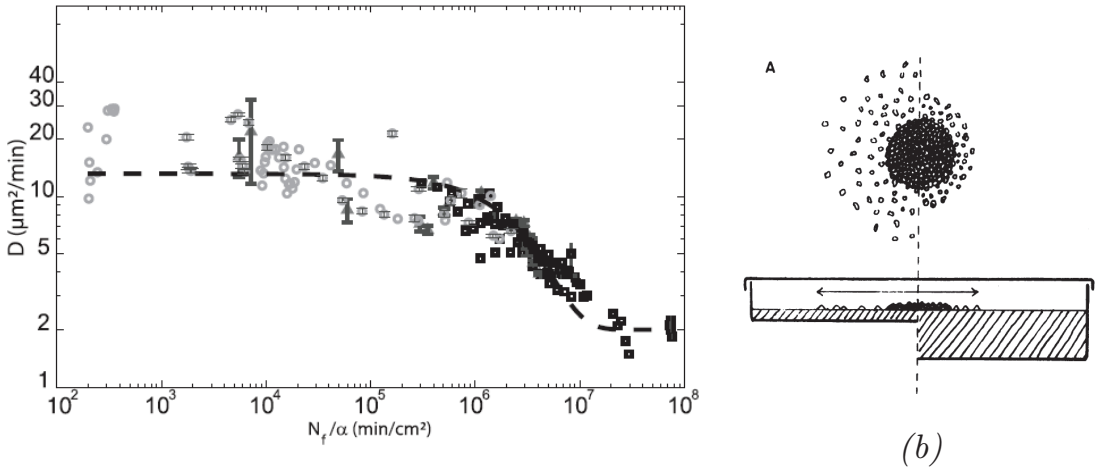


Figure 1.10: Summary of the cell polarity dynamics.
This cartoon illustrate how the cell polarity arises from the relative distribution of 'front inducers' (F, red) and 'back inducers' (B, green, left). F for instance promotes the protruding activity, while B promotes the contractile activity. It endows the cell with a polarity axis, which defines the direction of motion (right, modelling with a polar particle). The polarity axis may either change smoothly its direction (top), due to the fluctuations of the polarising factors, or be lost out of the blue, when the gradients of F or B destabilise suddenly (bottom). These phenomena give rise respectively to angular diffusion, where the turning angle probability depends on the sampling interval, or to tumbles, where the direction correlation is instantaneously relaxed.

4 Collective cell motion is not only the sum of single motions

As explained before, even though it is important to understand the phenomena at the single cell level, cells rarely perform their functions in complete isolation. The motility is no exception to this rule. On the contrary the cells often move with or in reaction to the presence of other cells, and the way they move in these situations is strongly affected by their interaction partners. We have made a distinction between interactions at a distance and contact interactions. In the following, we give an overview of these kinds of interactions in the particular case of motility, and of their possible outputs at both collective and individual levels.

4.1 Cell motility and the interactions



(a) (b)

Figure 1.11: Interacting at a distance.
(a) Dictyostelium cells secrete an unknown chemical factor that regulates its motility (from [18]). An empirical relation was found between the calculated concentration of this putative QSF and the diffusion constant of the cells. (b) Dictyostelium cells secrete a chemorepellent factor (from [19]). Sketch of the experiments: on shallow agar, the cells escape further than on deep agar, likely due to higher concentrations of this molecule.

Interactions at a distance. The most global and simple kind of interaction might be quorum-sensing, consisting in systems of secreted molecules to whose concentration the cells are sensitive. They can have many effects, from pathogenicity activation [20] to slowing down of proliferation [9], but their effects on the motility were only recently investigated [18]. In this last case, *Dictyostelium discoideum* cells have been shown to secrete an unknown QSF that acts to decrease the cells' diffusion constant in a concentration-dependent way (Fig. 1.11a).

A more complex and also more studied distant effect on motility is **chemotaxis**. It refers to the ability of cells to sense the gradient of a chemical and either move up (chemoattraction) or down (chemorepulsion). When the chemical is secreted by the cells themselves, it leads to a coupled dynamics that can be described in the scope of Keller-Segel models [21], and results in effective large-scale attraction or repulsion. The attractive case has been the most studied, and again pioneering works were made on *Dictyostelium discoideum*. In the well-known example of its developmental aggregation, upon nutrient starvation the cells start emitting periodically cyclic Adenosine MonoPhosphate (cAMP, a cyclic nucleotide). They are able to move up its gradient and to degrade it,

giving rise to spiral waves of concentration starting from an aggregation center and to cell streams towards this center. The functioning of chemotaxis has been studied at the molecular scale *e.g.* in *E. coli* [3], where it has been shown to rely on the regulation of the duration of run phases depending on the direction of motion.

In the reverse case, when cells move down the gradient of an endogenous messenger, the cells repulse each other. In *Dictyostelium*, such a mechanism was claimed to lead to a faster escape of the cells from spots made on shallow agar than from those made on thick agar plates [19] (Fig. 1.11b). More recently, the protein AprA was identified as an endogenous chemorepellent in *Dictyostelium* [22], and its mammal orthologue DP4 was found to be a chemorepellent for murine neutrophils [23].

Contact interactions. On the other hand, contact interactions, usually mediated by transmembrane receptors [24, 25], can also affect the migration properties of the cells. Mechanosensitivity is one of these mechanisms. The eukaryotic cells, especially in connective tissues as epithelia, are able to exert forces on their neighbour, which acts to direct the cell motion in the direction of lowest intracellular stress [26]. In some cases the forces can be transmitted by the substrate itself [27] but it is usually not the case in experiments on rigid glass or plastic dishes. This force transmission is clearly shown in ablation experiments: for instance, cells pulled by a leader at the edge of a growing epithelium retract upon release of the pulling force [28].

Even in the absence of strong adhesion and stress transmission, the collisions must have an effect on the motion, at least due to steric hindrance. This is particularly visible in the case of elongated cells, which undergo nematic alignment either due to hydrodynamic interactions in the case of swimming bacteria [29] or to real contacts for crawling bacteria [30] or mammal cells [31]. Beyond these passive effects, another mechanism, which was discovered more than 60 years ago [32], has gained renewed interest recently [33, 24]: it is called 'Contact Inhibition of Locomotion' – or CIL – and can be sketched as follows. When two cells encounter, they inhibit their protrusion in the contact zone, and repolarise in the direction opposite to the contact (Fig. 1.12). In this process the sensing is mediated by adhesion proteins, and especially cadherins, but also likely atypical cadherins, the Notch-Delta pathway, Ephrins or Nectins [24]. The intracellular signal is transmitted by the Wnt-planar polarity pathway and the small GTPases of the Rho family (RhoA, Rac, Cdc42...).

This last phenomenon has been more quantitatively studied in 1D since

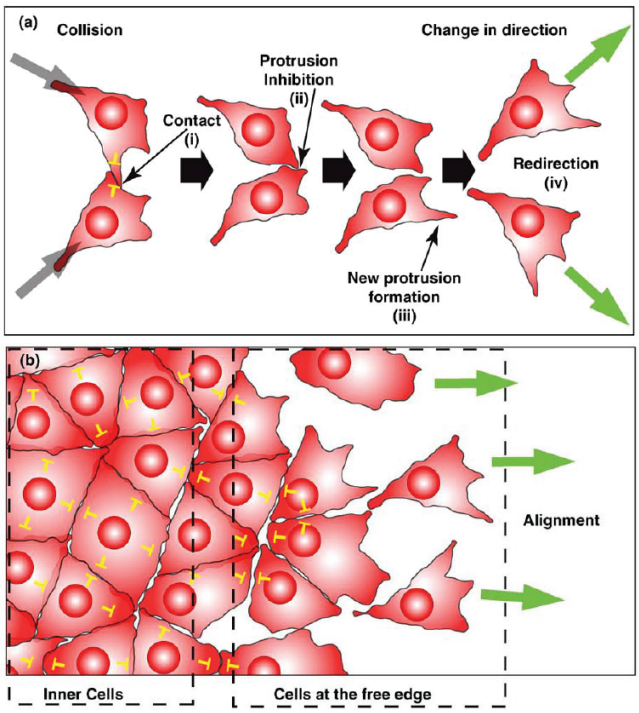


Figure 1.12: (a) Contact inhibition of locomotion in the case of two cells (from [24]). The protrusion are inhibited at the contact point and the cells repolarise away from each other. (b) In a cell layer, the inner cells are inhibited on all sides, while the cells at the free edge are polarised towards the empty space, allowing to form leader cells or to make room for the other ones.

the collisions are more frequent and easier to analyse quantitatively in this set-up [34, 35], but to our knowledge there is no quantitative study of the statistics of collisions in two dimensions.

4.2 The various effects of interactions

These cell-cell interactions are not neutral to the overall population dynamics. Actually, they may even lead to collective behaviours that are very different from those expected from single cells. Here, the structure of the population plays a lot on the observations. In particular, there is a clear distinction between effects in the core of a population, where the density is homogeneous in average, and on the edge, where gradients play a role.

In the core. We have evoked the existence of nematic interactions. In fact, alignment is a widely spread result of interactions. It manifests itself in spatial correlations: in bacterial systems [29, 30], but also in densely plated fibroblasts [31], nematic interactions can lead to orientational order over many cell lengths. Such long-range correlations also exist in migrating monolayers of epithelial cells [27, 36] or of fibroblasts [37, 38] (Fig. 1.13a). Generally speaking, these interactions act to increase the order and correlations. In the absence of alignment, however, the tendency is rather to a decrease, either through large-scale chemical interactions [18] or due to a collision-induced decrease of the

persistence [39]. In even denser connective cells, the tissue may exhibit a behaviour that reminds a glass transition, in which the cells are trapped by their neighbours [40, 41, 42].

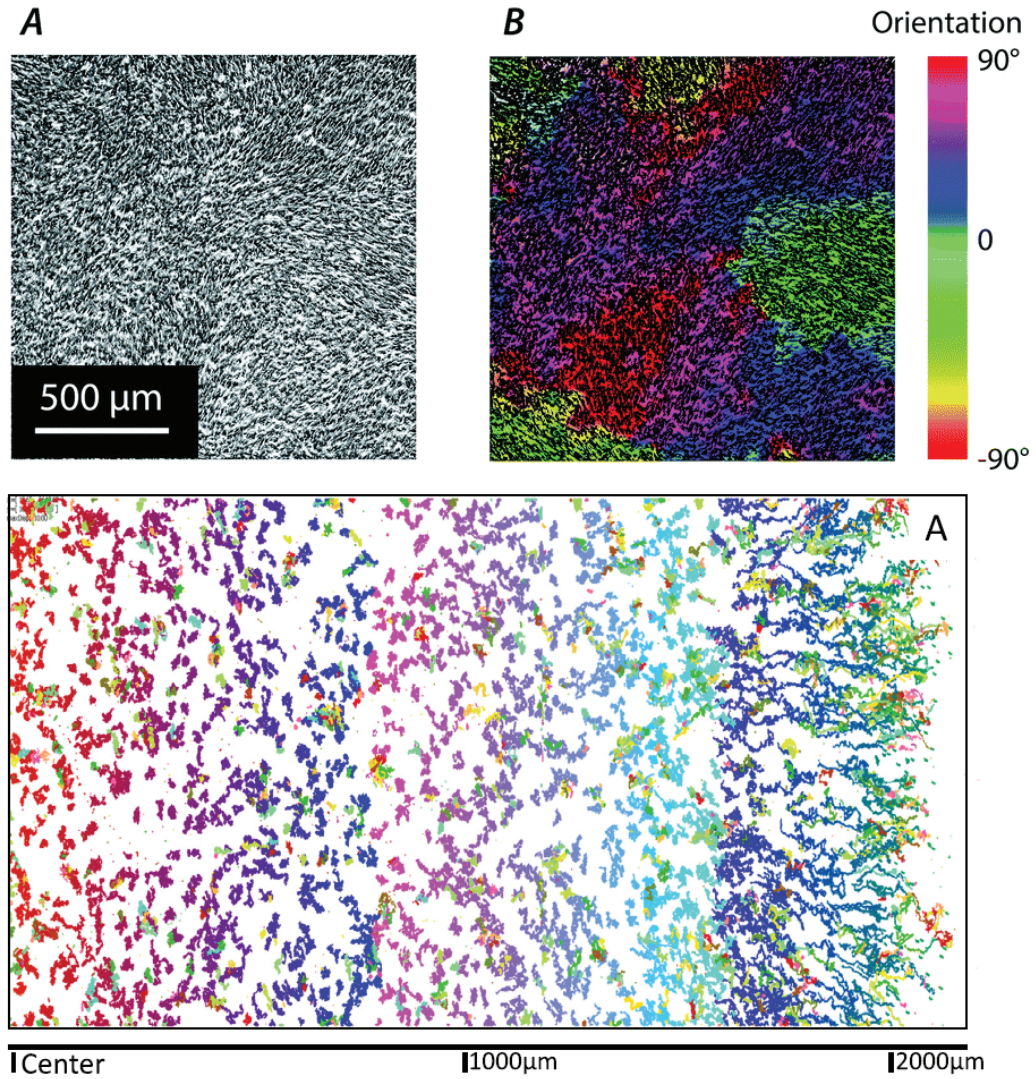


Figure 1.13: Effects of cell-cell interactions.
Top: In the core - local order in a high density layer of 3T3 fibroblasts (from [31]). Binarised field of view (left) and same image superimposed with the color-coded local orientation of the cells (right).
Bottom: On the edge - directionality is increased on the side of a monolayer. The color-labeled trajectories of epithelial cells are random and isotropic next to the center of the colony (left), while they are directed towards the empty space on the edge (right).

On the edge. A population of diffusing agents is expected to produce a net flux toward the void zones, because of the density gradient, but the motion should remain isotropic in the absence of external cues. It seems obvious that chemotaxis breaks this isotropy and creates a flux

of more or less directed cells. Actually, even contact effects can lead to such bias in the motion. Lange *et al.* [43] showed that cells on the edge of a monolayer are more directed towards the free space when they have more cell-cell adhesion (Fig. 1.13b). In the same scope, Nnetu *et al.* showed that this persistent motion of the interacting cells in the layer helped it to maintain its integrity, as the cells that escape will move more randomly, hence allowing the efficiently advancing sheet to catch them up. It is also much probable that the particular morphology of leader cells in expanding epithelia relies on the contact-free space asymmetry that they experience [28].

When the directionality is transmitted over the whole group, it gives rise to spontaneous collective motion, as it happens in small islands of adherent MDCK cells [44]. Concerning the effects clearly identified as CIL, they can improve the collective chemotactic response [45], and more unexpectedly lead to directed collective motion even when it is only associated with an attractive interaction [46].

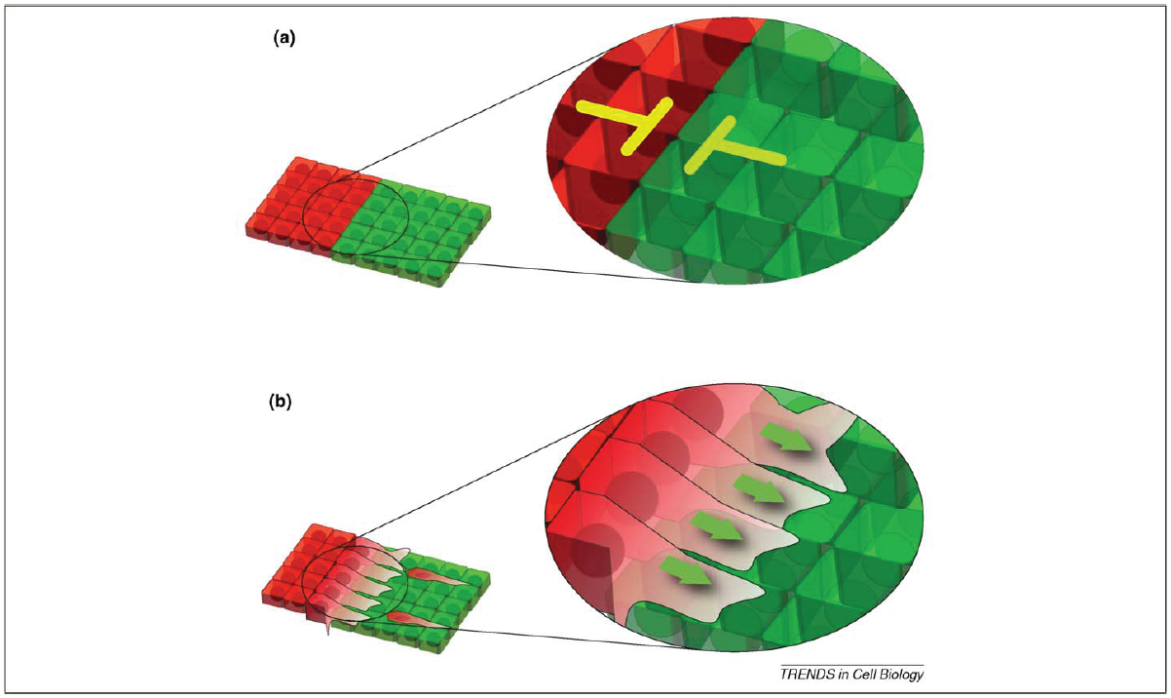


Figure 1.14: Interspecies contact inhibition of locomotion (from [24]). In the normal case (top), the motion is inhibited at the interface between two cell populations, preventing them from mixing. When interspecies CIL is inhibited for one of the tissues, its cells invade the other one.

The same kind of interaction also occurs between different populations of cells. The differences in interspecies contact inhibition of locomotion can promote the invasion of one tissue by another one [24] (Fig. 1.14). Conversely, the combination of chemotaxis of neural crest (NC) cells towards placodal cells combined with CIL between these two populations

leads to a chase-and-run behavior, where NC cells follow placodal cells that in turn escape, giving rise to a net collective displacement of both groups [47].

5 Active matter: a rising way of thinking collective effects

We have just seen that a great deal of interaction classes and related collective effects in the cell motility have been described, particularly in the last decade. Although the molecular mechanisms are in some cases quite well understood, the way by which interactions lead to ensemble behaviours remains usually elusive. Yet, these mechanisms of transition to a collective state have gained a huge interest in the community of theoretical physics, resulting in the appearance of a new field: active matter.

This name refers to systems of **self-propelled particles (SPPs)**, inspired mainly from living materials: molecular motors, cells, animals... Their particularity is that they consume energy taken from an external source, hence not being subjected to the laws of equilibrium thermodynamics. One of the aims of the discipline is to bring together very different systems, that share some common properties, and to understand how the collective dynamics arise from different classes of interactions.

The Vicsek model. The seminal paper of T. Vicsek [48] can probably be considered as the foundation stone of active matter. It considers point particles that move with a fixed speed along a diffusing direction – later termed active Brownian particles, or ABPs. These particles interact within a certain radius and have a tendency to align their direction along the average direction of motion of their neighbours. It was originally introduced to study, in the simplest way possible, the effects of alignment interactions in a group of moving agents.

It has been studied in depth, and recently many outstanding results were found [49, 50, 51]. In particular, it has been shown that a transition to order occurs when the density is increased or the noise decreased. This transition is similar to a liquid-gas-transition with a coexistence phase, where high-density ordered bands of defined size travel at a given velocity in a disordered gas phase (Fig. 1.15).

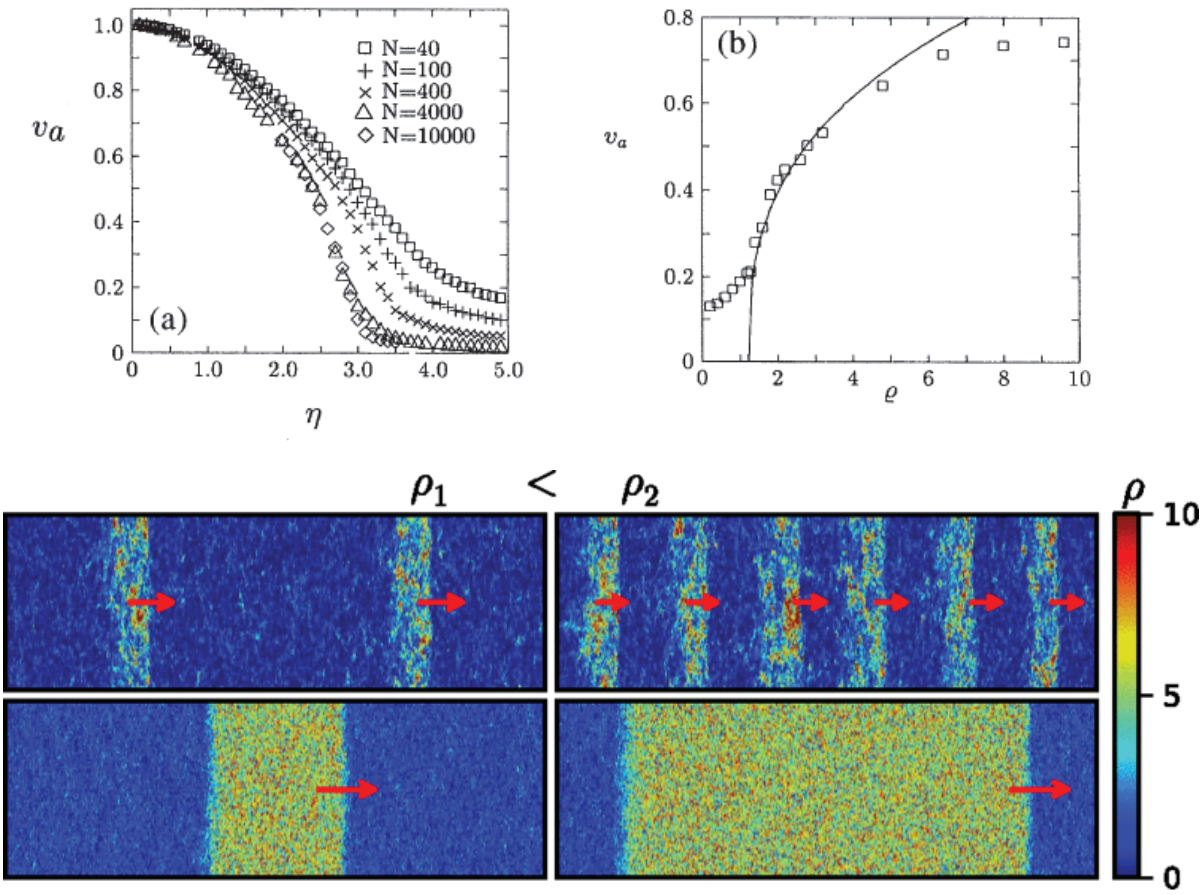


Figure 1.15: Transition to order in the Vicsek model.
Top: simulation results from [48]. When the noise amplitude η is decreased at constant density, or the density ρ increased at constant noise, the order parameter $v_a = \|\langle \mathbf{v} \rangle\|$ reaches non-zero values.
Bottom: simulation results from [50]. The transition to order leads to the formation of moving, ordered, high-density bands. When the density is increased, the band width remains constant and the number of bands increases in the Vicsek model (bottom panel, top), contrary to the simpler "active Ising model" where a single band grows (bottom).

Active Brownian particles and run-and-tumble particles. Despite its conceptual interest, the Vicsek model remains quite distant from real systems. More recently, people have started considering particles with a finite size, sometimes rod-shaped to mimic bacteria but most often hard disks. These particles move at a given speed, usually constant, along their polarity axis, and their rotational relaxation process defines the two basics classes of particles: **active Brownian particles, or ABPs** undergo smooth angular diffusion, while **run-and-tumble particles, or RTPs** go straight between sudden reorientation events.

These two different modes of relaxation can lead to similar dynamics to some extents, because both models belong to the bigger class of correlated random walks with exponentially decaying correlation function; yet, in some situations, they behave very differently. These two models have been thoroughly compared in a recent comprehensive paper [52].

Slowing down and clustering. The finite size of these particles implies the introduction of hard-core repulsion forces. In a collective context, these forces give rise to a pressure, which has been the focus of recent works [53, 54]. Another well-described effect of these forces is a density-driven decrease of the motility. Indeed, when such a particle faces an obstacle, it stops moving due to steric hindrance, and needs to wait for its direction to change enough to escape. It has been shown [55] that this effect leads to a decrease of the average speed that relates to the packing fraction ρ/ρ_m as $v(\rho) = v_0(1 - f\frac{\rho}{\rho_m})$, with $f \approx 0.9$. This effect is very important, for such a decreasing $v(\rho)$ or diffusion constant $D(\rho)$ can be responsible for **motility-induced phase separation**, or **MIPS**. This is a phenomenon of cluster formation without explicit attraction, driven by a positive feedback loop between motility and density: in suitable conditions, the natural fluctuations of density lead to a substantial local decrease of the motion; thus, the particles are trapped in high density zones, which in turn the faster particles coming from sparser areas can enter, hence accumulating. When the particles have no spatial extension, they eventually collapse into a single point, an effect that is counterbalanced in the case of hard disks.

The untapped potential of active matter. Active matter has got an important quality: it provides models that are at the same time simple, hence allowing to make general prediction on their basis, and realistic, which is essential to shed light on experimental results. Indeed, one of the problem of the theories of multicellular dynamics, and especially in the field of cancer, is that the proposed models are very specific, often lattice-based, and that their results themselves often depend on these specificities [56, 57, 58, 59]. Active matter avoids this pitfall, and in fact a few works have successfully used active matter modelling to get some insight in their experimental results [60, 61, 62, 42].

However, this potential remains, to my personal opinion, largely underexploited, and a lot could be learnt on the physics of such groups by easily implementable SPP simulations, or even mean-field or fluctuating hydrodynamics. Therefore, one of the ambitions of the present

manuscript is to contribute in setting active matter as the cornerstone for the description of collective behaviours arising from motile cells in interaction.

2. Methods for cell trajectory analysis

This chapter focuses on the object that served as a basis for almost all analyses in this thesis: the cell's trajectory. Firstly, we show the potential of this as a source of information on cellular processes. Secondly we review the classical models of cell trajectories Thirdly we describe the methods we used to obtain and analyse cell trajectory data; in particular, we give an overview of the typical properties of the motion of our *Dictyostelium* strain. Fourthly we calibrate our measurement tools against simulated data. The results and models contained in this chapter constitute the framework of analysis that we use in the following parts of this manuscript.

Traditionally, the trajectories of living organisms have been associated to breakthroughs in physics for long.
Here, the experimental observation of an apple's trajectory in the gravity field (center) served as the basis for Isaac Newton's theory of universal gravitation (1687). No doubt that he would also have agreed about the realism of hard spheres models for motile living matter (right). Gotlib ©DARGAUD, 2016.

Contents

1 Why do cell trajectories matter?

As it might have become apparent from **Chapter 1**, there are many scales at which cellular processes can be examined: at the subcellular level, active polymers under the control of regulation signals, produce the motion; at the cell scale, the cell deforms, moves, reacts to the environment; at the supra-cellular level, groups of cells exhibit coordinated, hindered or uncorrelated behaviours.

Diverse tools, including microscopy, allow measurements to be made at all these scales. Focusing on a particular cell, there is still a great deal of observations to be carried out: cell deformations, forces that it exerts or feels... This is very valuable information, but very complex ones at the same time. By zooming out a little while remaining at the single cell scale, one can then extract a single, simple piece of information: the cell's position.

Of course, defining the position is itself not so easy: should one take the center of mass of cell – or rather the barycentre of its projected area, since making a 3D reconstruction to measure this simple quantity would seem exaggerated? Or the centre of the nucleus? Or even, as some do [43], the position of nanoparticle internalised by the cell? The list of possible definitions is infinite, and we will see in **Section 3** that we used another one. Each definition yields a different result, but this can be overcome by considering the right scales.

A trajectory is a list of positions over time. This conceptual simplicity is an obvious strength, but beyond it, trajectories bear a lot of useful information on the system. Indeed, they are integrated manifestations of subcellular mechanisms and therefore they can provide insight into them: for instance the protrusion dynamics could possibly be inferred from the precise dynamics of *Dictyostelium* cells' trajectories [15, 63]. As the path followed by a cell also depends on its surroundings, it can inform either about the extracellular medium [64] or, as we claim in **Chapters 3** and **4**, about the interactions with other cells.

For a physicist, the interest of working with such an object is that it conveys notions that have been dealt with by statistical physics for decades. As such, trajectories can be subjected to – stochastic – kinetic models, which allows many analogies with well-understood systems. Similarly, when it comes to collective effects and despite some quite newly-studied specificities of active materials, statistical physics provides a toolbox that can be more or less directly applied to understand the underlying principles of observed phenomena. This last point is fundamental,

as this approach also enables to bring together systems whose particular internal functioning are distinct but whose motions are comparable, hence shedding light on the driving mechanisms that lead to similar population-scale behaviours.

2 Classical descriptions of cell trajectories

As it was mentioned in **Chapter 1**, the motility of a cell relies on stochastic processes. As a result, cell trajectories belong to the vast world of **random walks**. Note that we chose to use this expression to refer to actuals walks, *ie* with discrete steps, but also to the extended case of continuous-time trajectory whose direction changes smoothly (as for ABPs). In this section, we give an overview of this world, starting from the simpler discrete models. In particular we try to stress out the properties that are critical to a good understanding of cell motion and the special challenges raised by the distributed and intermingled nature of living matter.

2.1 Discrete random walks: from Simple Random Walk to correlated run-and-tumble dynamics

As we stated before, the world of random walks is vast, and the words of Jacques Monod about microbial growth (see **Chapter 1.2**) hold for this discipline. Moreover, the theoretical description of all existing kinds of random walks is outside the scope of this work. Yet, we will derive a few basic results in order to illustrate the governing principles of such walks. Note that unless specified otherwise, we consider particles moving at a constant speed v.

Simple Random Walk in 1D. The basic ideas of random walks are better understood in 1D. Consider a particle moving on a line with steps of duration τ and of length $a = v\tau$, with equal probabilities to go in both directions. Its displacement at step i is therefore $\delta x_i = \pm a$, with probability $p_\pm = 1/2$. It constitutes the mono-dimensional version of the **Simple Random Walk**, denoted **SRW**. It is a Markov process for the position, in so that the position at step i only depends on the position at step $i-1$; this kind of process is also called a position jump process. We are interested in the statistics of its displacement, namely its average position $\langle x \rangle$ and mean-squared displacement MSD $= \langle x^2 \rangle$ in time. At

time $t = n\tau$, the mean position is given by

$$\langle x \rangle = \sum_{i=1}^{n} (\frac{1}{2}a - \frac{1}{2}a) = 0, \tag{2.1}$$

and the MSD by:

$$\begin{aligned}
\langle x^2 \rangle &= \langle \left(\sum_{i=1}^{n} \delta x_i \right)^2 \rangle \\
&= \langle \sum_{i=1}^{n} \delta x_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} \delta x_i \delta x_j \rangle \\
&= \sum_{i=1}^{n} (\frac{1}{2}a^2 + \frac{1}{2}(-a)^2) \\
&= na^2 = v^2 \tau t,
\end{aligned} \tag{2.2}$$

because $\delta x_i \delta x_j = \delta_{ij} a^2$, where $\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$. This simple example illustrates that random walks are *diffusive* processes, for which the MSD is linear in time: $\langle x^2 \rangle \sim t$. We can also introduce the corresponding effective diffusion constant $D = \frac{v^2 \tau}{2}$ so that in this case $\langle x^2 \rangle = 2Dt$.

Correlation and the telegraph equation. Now, consider that what is subjected to fluctuations is not the position, but the velocity. In other words, $\overline{v} = \pm a/\tau$ can switch to the opposite value at each time step with rate λ (and so a probability $\lambda\tau$ at each time step). This is called a velocity jump process, and this introduces correlation, in so that the particle keeps a finite-time memory of its previous direction, hence the alternative name of **correlated random walk (CRW)**. Let us denote $\mathbb{P}^{\pm}(x,t)$ the probability that the particle is at position x at time t and moves with velocity $\pm a/\tau$ – then $\mathbb{P}(x,t) = \mathbb{P}^{+}(x,t) + \mathbb{P}^{-}(x,t)$ is the total probability to find the particle at position x at time t. An efficient way to treat this problem is to write the evolution equation for $\mathbb{P}^{+}$ and $\mathbb{P}^{-}$:

$$\mathbb{P}^{+}(x, t+\tau) = (1 - \lambda\tau)\mathbb{P}^{+}(x-a, t) + \lambda\tau \mathbb{P}^{-}(x-a, t) \tag{2.3}$$

$$\mathbb{P}^{-}(x, t+\tau) = \lambda\tau \mathbb{P}^{+}(x+a, t) + (1 - \lambda\tau)\mathbb{P}^{-}(x+a, t) \tag{2.4}$$

Expanding these in Taylor series and taking the limit a and $\tau \to 0$ while keeping $v = a/\tau$ constant yields

$$\partial_t \mathbb{P}^+ = -v\partial_x \mathbb{P}^+ + \lambda(\mathbb{P}^- - \mathbb{P}^+) \tag{2.5}$$
$$\partial_t \mathbb{P}^- = v\partial_x \mathbb{P}^- - \lambda(\mathbb{P}^- - \mathbb{P}^+) \tag{2.6}$$

After reasonable algebra, and using $\mathbb{P}(x,t) = \mathbb{P}^+(x,t) + \mathbb{P}^-(x,t)$, equations (2.5) and (2.6) yield the so-called telegraph equation:

$$\partial_t^2 \mathbb{P} + 2\lambda\partial_t \mathbb{P} = v^2 \partial_x^2 \mathbb{P} \tag{2.7}$$

The same approach could be used with the SRW, which would yield a diffusion equation. The telegraph equation (2.7) can be solved, but the solution is quite complicated [65]. However, one can compute the nth moment of $\mathbb{P}$ quite easily by multiplying equation (2.7) by x^n and integrating with respect to x. In particular, it yields again a zero net displacement, and the one-dimensional Fürth formula for the MSD:

$$\langle x^2 \rangle = 2D\left(t - \tau_p(1 - e^{-\frac{t}{\tau_p}})\right), \tag{2.8}$$

where $D = v^2\tau_p$, and $\tau_p = \frac{1}{2\lambda}$ defines the persistence time of the movement. We can introduce the velocity autocorrelation function, which measures the memory of the direction over a time interval δt:

$$C(\delta t) = \langle \overline{v}(t)\overline{v}(t+\delta t) \rangle \tag{2.9}$$

Here, we use the total (integrated in space) probability that the particle goes to the right $p_r(t)$ at time t. The time derivative of p_r writes

$$\dot{p_r} = -\lambda p_r + (1-\lambda)p_r = (1-2\lambda)p_r, \tag{2.10}$$

which is obtained in a similar way as equations (2.5) and (2.6). There are four cases to average over in the right-hand side of equation (2.9): the particle goes right at t, then right at $t + \delta t$, or left, then left; right, then left, or left, then right. The product will equal v^2 in the two former cases, $-v^2$, in the two latter. Therefore, $C(\delta t)$ writes:

$$\begin{aligned} C(\delta t) =& v^2[p_r(t)p_r(t+\delta t) + (1-p_r(t))\,(1-p_r(t+\delta t)) \\ & - p_r(t)\,(1-p_r(t+\delta t)) - (1-p_r(t))\,p_r(t+\delta t)] \\ =& v^2\,[1-2p_r(t)-2p_r(t+\delta t)+4p_r(t)p_r(t+\delta t)] \end{aligned} \tag{2.11}$$

Differentiating with respect to δt:

$$\begin{aligned} \partial_{\delta t} C &= -2v^2\partial_{\delta t}p_r(t+\delta t) + 4v^2 p_r(t)\partial_{\delta t}p_r(t+\delta t) \\ &= -2\lambda v^2\,[1-2p_r(t)-2p_r(t+\delta t)+4\lambda p_r(t)p_r(t+\delta t)] \\ &= -2\lambda C \end{aligned} \tag{2.12}$$

Assuming p_r is in the stationary regime, C does not depend on t but only on δt. Then it comes

$$C(\delta t) = e^{-2\lambda\delta t} = e^{-\frac{t}{\tau_p}} \tag{2.13}$$

This concept of exponentially decreasing correlation function is a fundamental property of many random paths. Note that the MSD can be obtained by integration of the correlation function. In particular, for an exponentially decaying correlation function, one recovers Fürth's formula for the MSD.

These results can be extended to higher dimensions, in particular in 2D which is our focus, either on or off lattice. In particular, the 2D off-lattice version with no correlation in the turning angle is the aforementioned run-and-tumble particle model.

The general continuous time CRW adds a possible correlation in the turning angles: in this model, the particle moves along direction $\mathbf{e}_\theta$ at constant speed during a time drawn from an exponential distribution with characteristic time τ, and then tumble instantaneously with a turning angle $\delta\theta$ drawn from a distribution that may not be flat – *e.g.* the von Mises distribution, or the wrapped normal or Cauchy distribution [65]. It just changes the mathematical details, but not the general behaviour that:

(i) the autocorrelation function decays exponentially: $C(\delta t) \sim \exp(-\frac{\delta t}{\tau_p})$ where $\tau_p = \tau$ in the case of run-and-tumble particles (RTP, see **Chapter 1**), but not necessarily in other situations;

(ii) at long times $t \gg \tau_p$, the motion is diffusive with $\langle x^2 \rangle = 4Dt$ in 2D;

(iii) the MSD is given by the Fürth formula;

(iv) the diffusion constant writes $D = \frac{v^2 \tau_p}{2}$ in 2D,

where *(i)* implies *(ii)–(iv)*.

2.2 Smooth trajectories

Such properties hold in the case of smoothly turning trajectories, provided the correlation function remains exponential. Two particular cases are interesting to the extent that they are both quite simple and widely studied.

The Ornstein-Uhlenbeck (OU) process has been used to model cell motion since the 1970s [66, 67]. Originally proposed to describe Brownian motion [17], it describes the velocity's dynamics as follows:

$$\begin{cases} \dot{\mathbf{r}} = \mathbf{v} \\ P\dot{\mathbf{v}} = -\mathbf{v} + \sqrt{2D_v}\,\boldsymbol{\eta} \end{cases}$$

It can be interpreted as a Langevin equation with a damping of magnitude $1/P$ and a Langevin force of magnitude $\sqrt{2D_v}/P$, $\boldsymbol{\eta}$ being a Gaussian white noise with unit variance. This model admits a Gaussian stationary distribution for v with a variance that tends to PD_v. The velocity autocorrelation function decays exponentially with relaxation time P and the MSD obeys Fürth's formula. It has been very widely used to describe the trajectories of isolated cells, and even recently integrated in a model of interacting cells [61] which matches very satisfactorily experimental data obtained from epithelial monolayers.

However, the meaning of its terms in the context of crawling cell migration is fuzzy: one expects rather over-damped dynamics than inertial dynamics with friction. Also, here the stochasticity comes in the form of a Langevin force, while in cells it rather occurs from intracellular processes. As such, despite its descriptive power, the OU process remains mainly phenomenological.

Another simple model that in my opinion deserves attention is the active Borwnian particle (ABP) model. It assumes that the particle moves with constant speed along a polarity axis, which undergoes angular diffusion:

$$\begin{cases} \dot{\mathbf{r}} = v\mathbf{e}_\theta \\ \dot{\theta} = \sqrt{2D_r}\eta \end{cases} \tag{2.14}$$

where η is a Gaussian white noise with unit variance [52]. It decouples the internal dynamics that governs the cell polarity from the motion itself, even though a $D_r(v)$ relation could be considered, as suggested by Maiuri *et al.* [16].

A mixture of ABP and RTP provides a comprehensive model for a self-propelled particle (SPP) with constant speed, which is better described by its master equation – because the presence of instantaneous reorientations, Langevin equations are not practical – namely

$$\begin{aligned} \dot{\mathbb{P}}(\mathbf{r}, \mathbf{e}_\theta, t) = &- \boldsymbol{\nabla} \cdot [v\mathbf{e}_\theta \mathbb{P}(\mathbf{r}, \mathbf{e}_\theta, t)] + \boldsymbol{\nabla}_{\mathbf{e}_\theta} \cdot [D_r \boldsymbol{\nabla}_{\mathbf{e}_\theta} \mathbb{P}(\mathbf{r}, \mathbf{e}_\theta, t)] \\ &- \omega \mathbb{P}(\mathbf{r}, \mathbf{e}_\theta, t) + \frac{\omega}{2\pi} \int_{\theta'=-\pi}^{\pi} \mathbb{P}(\mathbf{r}, \mathbf{e}_{\theta'}, t) d\theta' \end{aligned} \tag{2.15}$$

Here, $\mathbb{P}(\mathbf{r}, \mathbf{e}_\theta, t)$ is the probability to find the particle at position $\mathbf{r}$, with orientation $\mathbf{e}_\theta$, at time t. On the right-hand side, the first term is the advection due to the self-propulsion (we forget translational diffusion which is less relevant for a migrating cell); the $\boldsymbol{\nabla}_{\mathbf{e}_\theta}$ term denotes the angular diffusion; the first term on the second line accounts for the loss of particles initially oriented along $\mathbf{e}_\theta$, by tumbling at a rate ω, and the last term is the corresponding gain.

All these models exhibit some differences in their microscopic definition, and especially in the way the directional persistence relaxes. Although these discrepancies can show up on large-scale observations in certain situations [52], the averaged properties of trajectories under no external constraint are essentially similar. In particular all these models exhibit an exponentially decaying velocity autocorrelation function with a single relaxation time – which can be a combination of several times, as in the last one: $\tau_p = \frac{1}{D_r+\omega}$.

2.3 Further than Fürth

This matches very well the experimental measurements on cell trajectories in many cases [66, 67], but more recent works pointed out divergences from the expected properties of such models [68, 69, 70, 71, 18]. In particular, some revealed the existence of *two* different relaxation times in

the velocity autocorrelation. It could arise directly from the existence of two subpopulations of cells, but this hypothesis is usually ruled out quite right up. There are – at least – two other ways to account for this double relaxation.

The devil is in the details. The approach introduced by Selmeczi *et al.* [68] consists in building a theory from experimental data, based on a general Langevin-type equation for the velocity:

$$\dot{\mathbf{v}} = -K * \mathbf{v} + \bar{\bar{\boldsymbol{\sigma}}}\,\boldsymbol{\eta} \tag{2.16}$$

where $\boldsymbol{\eta}$ is a Gaussian vectorial white noise as usual. K is a kernel that can contain memory and can depend on v, and $\bar{\bar{\boldsymbol{\sigma}}}$ is an order two tensor to account for a potential noise anisotropy. It can depend on v as well.

They use a complete dataset to infer the form of K and $\bar{\bar{\boldsymbol{\sigma}}}$. In particular, the mathematical form of the velocity autocorrelation function sets constraints on the form of K, and the properties of the distributions of $\dot{\mathbf{v}}$ for different v inform about the symmetries and functional dependencies of K and $\bar{\bar{\boldsymbol{\sigma}}}$.

In that fashion, they derive the simplest possible model that fits the data well. In particular, the two correlation times arise from a complex interplay of at least the "friction coefficients" and the memory kernel's time scale. This method is interesting in that it is solely data-based, and it provides an accurate description that seems cell-type specific [68, 63]. Then the form of K and $\bar{\bar{\boldsymbol{\sigma}}}$ may inform on the dynamics of the cell polarity and propulsion. Yet, the interpretation of the various terms and parameters is not so straightforward. Moreover, this data driven procedure does not exclude other types of models.

Multimodality in cell motion. In the previous models, the multi-relaxation behaviour arose from entangled processes acting on different time scales. This is also the case in a recent model [61], developed in the purpose of modelling interacting cells, that relies on an Ornstein-Uhlenbeck process whose noise is itself correlated through another Ornstein-Uhlenbeck process. Conversely, in the last ten years, there has been a rising opinion that crawling cells could exhibit multimodal motion, with an alternation of steps (also "portions of trajectory" hereafter) with a single relaxation time.

This property seems obvious for instance in the case of swimming bacteria, with a well-described alternation of running and tumbling phases in *E. coli* for instance [72]. There has been evidence that such bimodality could occur in the motion of mammalian cells [71, 73] or *Dictyostelium* cells [18, 74, 75] as well. Theoretical models of bimodal motion have just started being studied extensively [76, 77]. One should distinguish between two cases, with respect to the distribution step durations in each in mode. When both distributions are exponential, *ie* the switch processes are Poissonian, one can compute exact results, especially for the MSD [76]. When the modes at stake consist in two rotational diffusions with different persistence times, one can even express the velocity autocorrelation function as:

$$C(t) = Ce^{-\gamma t} = C'e^{-\gamma' t} \tag{2.17}$$

where C, C', γ and γ' are functions of d_1, d_2, λ_1 and λ_2, respectively the rotational diffusion constants and inverse means of the exponential distributions of the two modes (see **Appendix A** for detailed expressions). The limit $d_i \to 0$ (resp. $d_i \to +\infty$) represents a completely random (resp. ballistic) motion for mode i. In particular, $C, C' \geqslant 0$, $C + C' = 1$, $d_1 \leqslant \gamma' \leqslant d_2 < \gamma$.

When the step-size distribution are non-Poissonian, it is more difficult to obtain analytical results, but it is still possible to compute the propagator with a somehow symbolic method, and to apply this to specific cases [77].

Distributed, non-constant parameters. The reader would have noticed that to this point the picture has become more and more fuzzy. However, the considered models still have strong assumptions. In particular, even if several modes of motion or of directional relaxation are allowed, until here the underlying parameters have been kept constant at fixed values. Actually, it could very well be that these parameters are not constant.

For instance, there is no example of crawling cell trajectories without fluctuations in the speed. The Ornstein-Uhlenbeck-Selmeczi-like models exhibit this property, because the noise acts on the whole velocity vector. On the contrary, in the SPP models where the noise acts on the polarity axis independently from the motion itself, fluctuations in the speed must be taken explicitly into account. This approach is therefore relevant to situations in which the fluctuations in direction and in speed originate

in different mechanisms [78]. Given the complexity of motile cells, this does not seem absurd, even though they could also well be correlated in some way [16]. In case speed and direction are independent, Peruani and Morelli [78] showed that, if the speed correlation decays exponentially with constant λ_v and the direction diffuses with diffusion constant λ_θ, the velocity autocorrelation function writes

$$\langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle = \langle v \rangle^2 e^{-\lambda_v t} + (\langle v^2 \rangle - \langle v \rangle^2) e^{-(\lambda_v + \lambda_\theta)t} \tag{2.18}$$

In addition, the persistence itself could evolve in time, either by discrete switches as in bimodal motion, or in a continuous way, because the persistence itself relies on an underlying stochastic process. Among the possible causes for that, the substrate may not be completely homogeneous, which is quite true when it comes to collagen or fibronectin-coated surfaces for instance; alternatively, the cells can undergo bimodal motion with spread distribution of persistence time in each mode; even more simply that time, the persistence can be distributed over the population, but contrary to speed it is hardly measurable without very long trajectories.

2.4 Summary

The attribute *vast* used to qualify the world of random walks having probably become meaningful to the reader, it might make sense to sum up all the above in a compact way.

Firstly, it is important to underline that on sufficiently long time scales, *every* isotropic random walk becomes **diffusive** (except in very particular, mostly theoretical cases). All the so-called "anomalous" dynamics that could be described arise from too short observation windows.

Secondly, at shorter time scales, the dynamics can be varied. Many models fall in Fürth's class: their velocity autocorrelation function (VACF) decays exponentially with a **single relaxation time**. However, they can have either a constant (*e.g.* ABPs, RTPs, CRW) or distributed speed (Rayleigh-distributed in the case of the OU process).

By contrast, many models can account for a sometimes observed **double exponential decay** of the VACF. The speed distribution can again vary from a Dirac to a spread one. The two relaxation times can either arise from the noise, which is itself correlated, from a memory "frictional" kernel, or from explicit bimodal motion.

Thirdly, we distinguish between two main means, by which the noise is accounted for: in **OU-like dynamics**, it acts directly on **v** in a non-overdamped equation; in **SPP models**, it acts on the polarity angle, hence decoupling the latter from the speed dynamics, and the equation of motion is usually overdamped.

Understanding the motion requires to be able to find a good model for the trajectories and to interpret its parameters. At the other end of the spectrum, deciphering the role of cell-cell interactions also demands to first have a good model of single cell motion, essentially to measure the changes induced by the interactions. From this perspective, it is important to get, first, cell trajectories, and then reliable – and the less arbitrary possible – tools to measure properties from these data.

3 Obtaining and analysing cell trajectory data

In this section we describe the way we obtain cell trajectory data, by tracking cells on time-lapse microscopy images. Then we present the main analysis tools that we use to characterise the trajectories. Last, we simulate random paths that exhibit similar expected properties as the experimental ones, and we test our analysis procedures on them to estimate their accuracy and especially how the results are affected by the discrete nature of the data.

3.1 Imaging and image processing

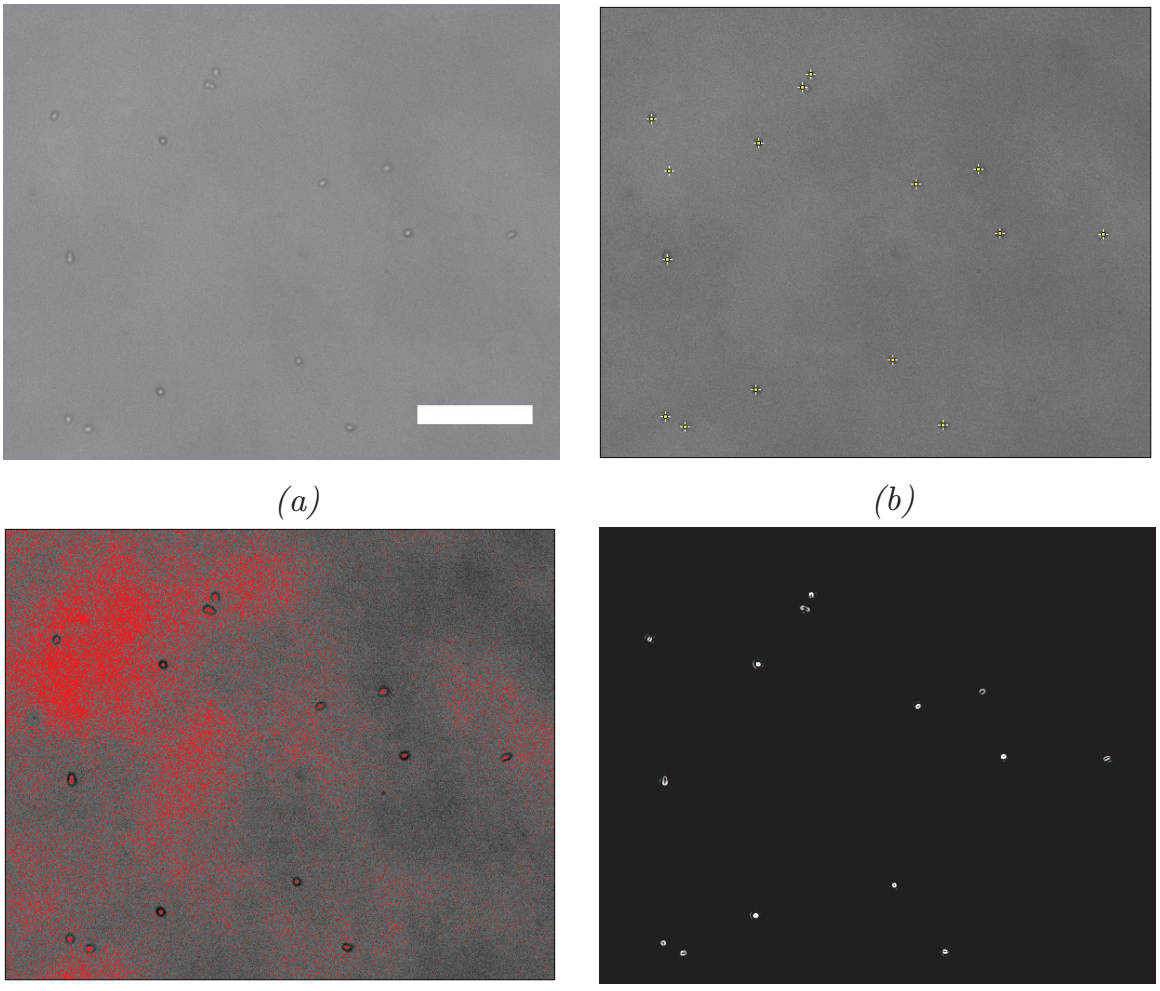
Unless otherwise mentioned, all the movies were taken with the following specifications:

– **material:** Nikon TE2000 microscope with automated XY-stage, Andor Zyla camera (field width: 2160x2560 pixels, pixel size: 6.5 µm); the pictures in **Chapter 3** were made at 4X (1.63 µm/pixel) or 3X (2.17 µm/pixel) magnification, while those of **Chapter 4** were taken at 10X magnification (0.65 µm/pixel), in slightly defocused bright-field (for the cells to appear as very bright spots, see Fig. 2.1a);

– **temperature control:** to keep the temperature constant at 22°C, a home-made box of polystyrene was built to contain the whole microscope; a heat-exchanger containing circulating water was placed in the box; the water temperature was adjusted to keep the air temperature constant thanks to a water bath (Julabo, Germany) connected to a P1000 temperature probe;

– **time-lapse:** time-lapse images were recorded automatically using MicroManager software, with a typical 15 ms exposure-time and a various time delays: $\delta t \simeq 1$ min in **Chapter 3** and $\delta t = 20$ s in **Chapter 4**.

To detect the cells' position, three different techniques could be used, all run by custom ImageJ macros, taking advantage from the brightness of the cells (Fig. 2.1).



(a) *(b)* *(c)* *(d)*

Figure 2.1: Cell detection.
(a) Typical picture obtained by means of low-magnification defocused microscopy. Scale bar: 200 µm. *(b) Detection using the Find Maxima function. The plus signs denote the detected maxima. (c) Thresholding. The red areas correspond to zones above the threshold before binarisation. This picture does not seem to be really suitable for this kind of detection. (d) Edge detection. The bright zones denote detected edges. Further post-processing is still required to segment the cells.*

Find Maxima. This technique used the "Find Maxima" built-in function to automatically detect the local light maxima on the picture (Fig. 2.1b).

It is very simple to use since it only demands defining the correct 'noise tolerance' parameter before running the macro. The maximum of light is located at the highest point of the cell, which correlates nicely with the location of the nucleus, and is very close to the center of mass of the cell since *Dictyostelium* cells are quite round-shaped.

Yet, its outputs are only pixel-resolved, which may cause some problems for the analysis. Furthermore, polarised, elongated cells often exhibit two local maxima which in some cases were identified as two different cells, although this last issue was secondary.

Threshold. Since the cells are very bright, in some cases a simple thresholding of the picture is enough to detect the cell contour with a good accuracy (Fig. 2.1c).

The drawback of this technique is that in dense cell populations, it may be unable to separate touching cells. Increasing the threshold value can help to this: then the cell area may be underestimated, but this is not a serious issue as anyway the cell contour is not accurate because of the out-of-focus nature of the pictures. In some cases, this technique may lead to elongated cells being cut in two parts, as with the Find maxima technique.

Edge detection. This technique uses the "Find Edges" ImageJ built-in function. It is a variant of the previous one since the aim is also to detect the cell contours in order to place the center of mass with sub-pixel resolution (Fig. 2.1d).

Depending on the pictures, one or the other technique can be used. The accuracy of the detection was controlled each time manually on a few frames over the movie duration.

Of course, such automated analysis over long-time movies supposes that the pictures remain similar throughout the movie, so that the parameters – noise tolerance or threshold values – hold for all pictures, and errors in this very first step of the data treatment occur with a non-negligible rate, from 1% to 10% in some difficult cases. On the other hand, it would not be possible to detect enough cells for a meaningful analysis manually, so this drawback has to be accepted with humility.

Tracking and data structure. After this detection step, the centres of mass are stored in a 3-columns table containing x-y coordinates and frame label for each position. Then the **track.m** Matlab code (J. Crocker, http://www.physics.emory.edu/faculty/weeks//idl/) is used to link the positions from subsequent frames. Basically, this code uses the least squares technique, *ie* it minimizes the squared distance between couples of positions from subsequent frames, over all allowed combinations (which number is limited thanks to the 'Maxdisp' parameter that sets the maximum distance allowed) to find the most likely matches. A fourth column is added to the table, containing the cell's label.

Sometimes the cell is lost, leading to an exponential distribution of trajectory lengths (N_i). Then these gathered trajectories can be treated to measure various single-cell or averaged parameters.

3.2 Measured quantities.

Velocity. The velocity is defined as $\mathbf{v} = \dot{\mathbf{r}}$. We compute it from our discrete data as $\mathbf{v}_i = (\mathbf{r}_{i+1} - \mathbf{r}_i)/\delta t$, where δt^{-1} is the frame rate, and i is the frame label. The speed is directly $\|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2}$.

Mean-squared displacement. The MSD is computed in a similar way (for each cell) as $\mathrm{MSD}(n\delta t) = \frac{1}{N-n}\sum_{i=1}^{N-n} \|\mathbf{r}_{i+n} - \mathbf{r}_i\|$, where N denotes the length of the considered piece of trajectory.

Diffusion constant. To compute the diffusion constant D of *a population*, we first keep only the trajectories whose lengths are larger or equal to N (generally so that $N\delta t \simeq 90\,\mathrm{min}$, see below). Then we compute the MSD at $t = |N/3|\delta t$, and $D = \langle \frac{\mathrm{MSD}(t)}{4t} \rangle$, where the angle brackets denote the population averaging. We take $t \geqslant 15$ min [79]; typically $t = 30$ min in **Chapter 3** and $t = 20$ min in **Chapter 4**.

Velocity autocorrelation function. The VACF is computed in the direct space as the biased estimator $C_{\mathbf{v}}(n\delta t) = \langle \frac{1}{N-n}\sum_{i=1}^{N-n} \mathbf{v}_{i+n} \cdot \mathbf{v}_i \rangle$. We also use most of the time the velocity *direction* autocorrelation function (VDACF) $C(t)$ obtained by replacing any $\mathbf{v}_i$ by $\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$ in the above formula. The relation $C_{\mathbf{v}}(t) = \langle \|\mathbf{v}\| \rangle C(t)$ is well verified in our data.

Coefficient of movement efficiency. The CME provides an efficient way to estimate the amount of directionality at a given time scale, without the need for huge statistics over a long time window, as the VACF. It is defined as the end-point displacement of a cell in the interval $n\delta t$, divided by the total path length during the same time,

$$CME_\tau(t) = \frac{\|\mathbf{r}(t+\frac{\tau}{2}) - \mathbf{r}(t-\frac{\tau}{2})\|}{\int_{t'=t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} \|\mathbf{v}(t')\| dt'}, \tag{2.19}$$

or in its discrete version (measured at time $t = i\delta t$, and over a time window $\tau = n\delta t$):

$$CME_i^{n\delta t} = \frac{\|\mathbf{r}_{i+n} - \mathbf{r}_i\|}{\sum_{j=i}^{n-1} \|\mathbf{r}_{j+1} - \mathbf{r}_j\|} \tag{2.20}$$

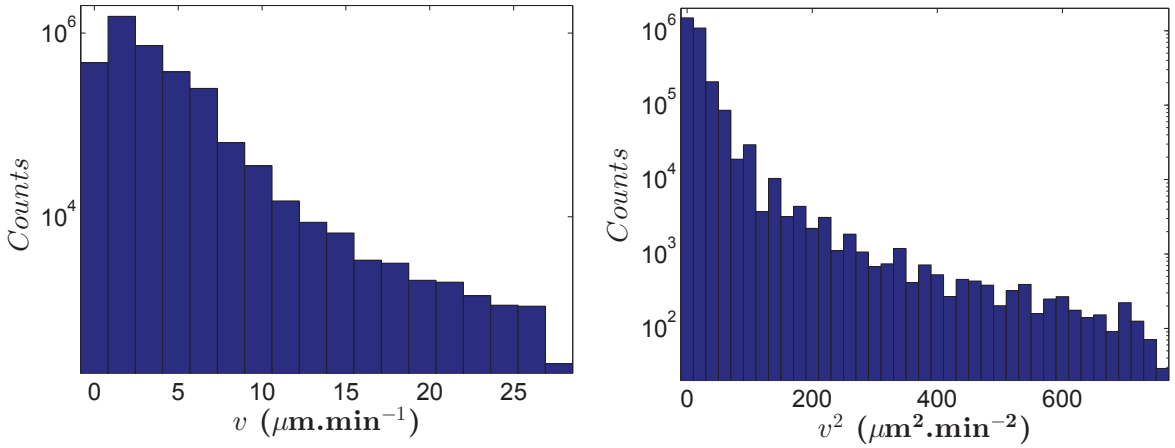
Of course, other quantities are measured in the course of this manuscript, but the above are the main one used to characterise the trajectories.

3.3 Typical properties of *Dictyostelium*'s trajectories

In this section we report the main characteristics of the trajectories of isolated *Dictyostelium* cells from the AX2 strain, which is the 'eldest' of axenic strains (it was derived from AX1, the – lost – first *Dictyostelium discoideum* cells grown without bacteria). The results given here are from 156 independent experiments made on dilute populations of cells in fresh HL5 medium, with a total of about 25,000 full length ($\geqslant$ 90 min long) trajectories and $\sim 3 \times 10^6$ single data points (speed - acceleration *etc...*).

The detection method was the "Find Maxima" one, as it is the standard in this work, with $\delta x = 1.63$ µm/pixel and a $\delta t = 1$ min time lapse.

Speed. First, note that given the pixel-resolved nature of the positions, the detected speeds take well defined discrete values, the first ones being $\{1; \sqrt{2}; 2; \sqrt{3}; 2\sqrt{2}; 3; \sqrt{10}...\} \times \delta v$, where $\delta v = \frac{\delta x}{\delta t} = 1.63$ µm.min^{-1}. The speed distribution is shown in Figure (2.2). Its mean is $\langle v \rangle = 3.0$ µm.min^{-1}, and its standard deviation $\sigma_v = 2.5$ µm.min^{-1}. The distribution is far from Gaussian, Rayleigh or log-normal. It is not exponential either, even though it has a similar behaviour over a certain range of intermediate speeds.



(a) *(b)*

Figure 2.2: Speed distribution.
(a) Logarithmic plot of the speed histogram. It is somehow straight from the second bar to $v \sim 15$ µm.min^{-1}, exhibiting a behaviour close to an exponential decay. (b) Logarithmic plot of the histogram of squared speed. It shows that the distribution is not Gaussian, which would produce of linear decrease.

Mean-squared displacement and effective diffusion constant. Taking the trajectories longer than 90 min, we then plotted the MSD against the time lag Δt. As reported in Ref. [18], the MSDs of individual cells are very spread (Fig. 2.3a). Its distribution at large time looks also log-normal (Fig. 2.3d), consistently with these previous observations [79]. The exponent of the average MSD is about 1.4, showing that the observation time is around the transition from ballistic to diffusive behaviour.

In linear scale, the MSD seems to exhibit a linear behaviour after a transient 'acceleration' phase (Fig. 2.3b). From an affine fit at $\Delta t = 20 - 30$ min, we extract the diffusion constant $D_{fit} = 13.1 \pm 0.2$ µm^2.min^{-1} (95% confidence interval). Then, we plotted MSD/Δt against the time lag. This plot is customary in the team as it shows the diffusive character through the plateau that emerges at $\Delta t \sim 10\tau_p$, and it makes the discrepancies from OU dynamics at short times clear. Here, the maximum Δt is still too short to see really the plateau. However, we can still define an apparent diffusion constant:

$$D = \frac{\mathrm{MSD}(\Delta t_{max})}{4\Delta t_{max}} \tag{2.21}$$

Here, $D = 10.9$ µm^2.min^{-1}. Of course, this measurement induces an

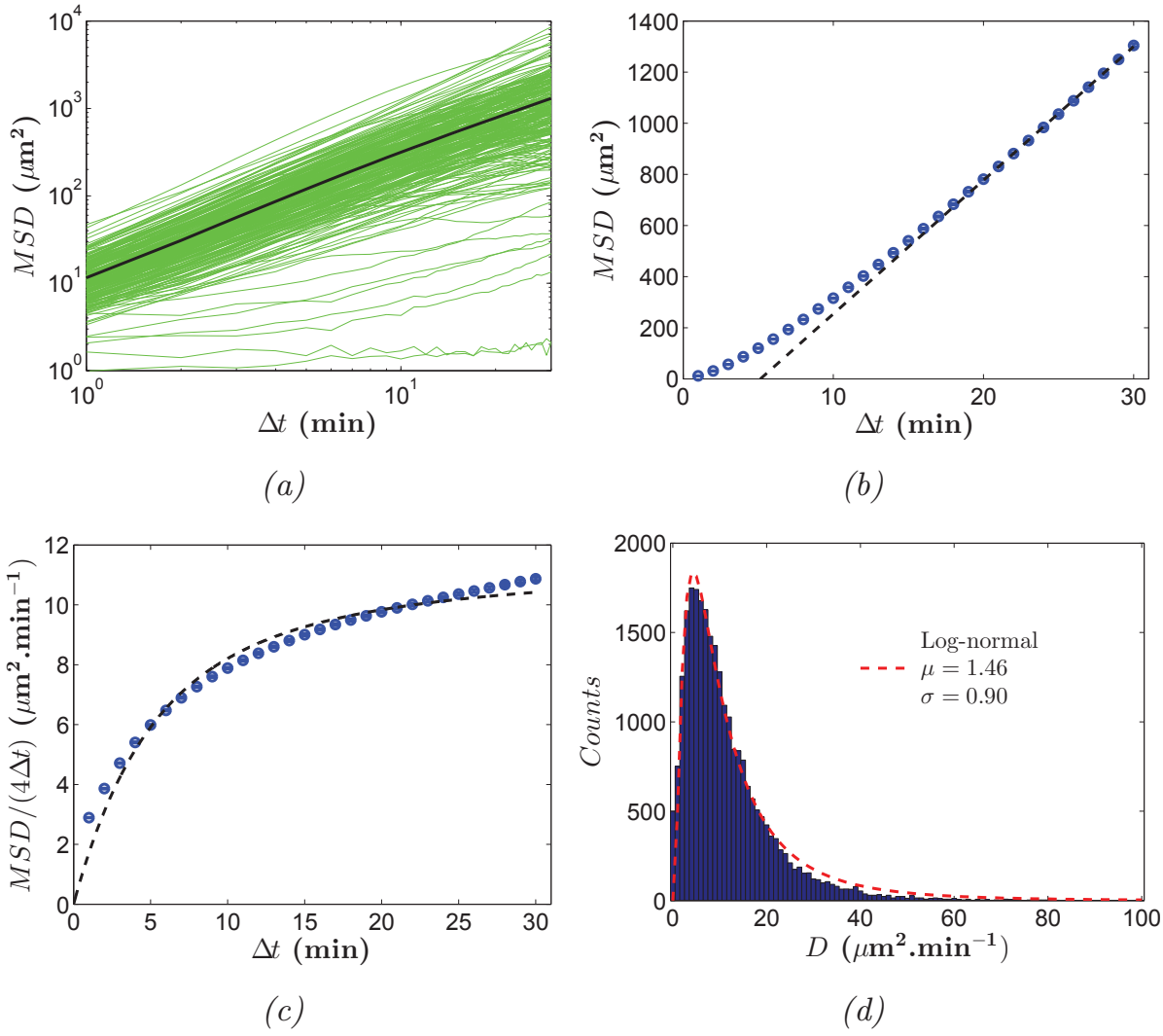


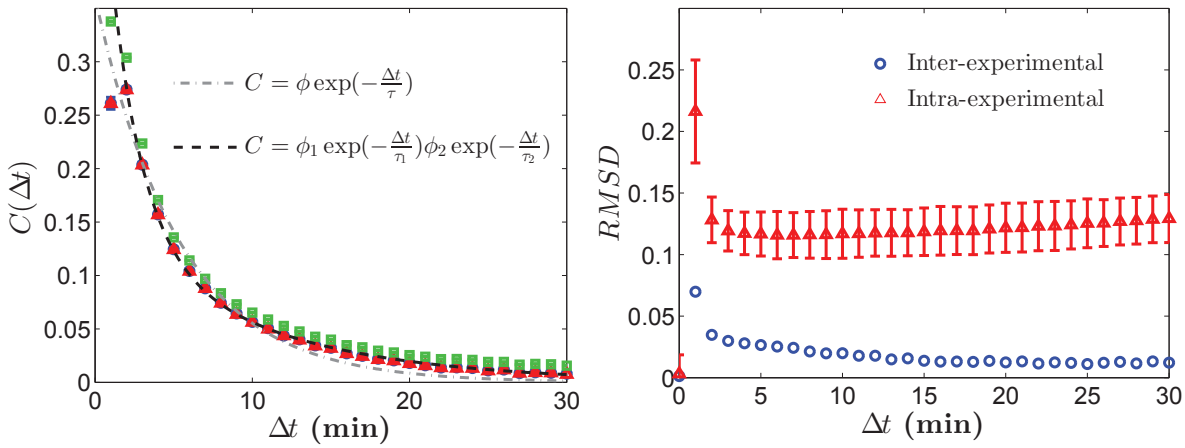
Figure 2.3: Mean-squared displacements (MSD)
(a) Log-log plot of individual MSDs (green, thin lines), and average (black thick line). For the sake of clarity, only 1% of the individual trajectories, chosen at random, are shown. The average has a slope of ~ 1.4. *(b) Averaged MSD in linear scale. The dashed line represents a fit of the linear part of the curve:* $MSD = 4D(t - t*)$ *with* $D = 13.1 \pm 0.2\ \mu\text{m}^2.\text{min}^{-1}$ *(best fit and 95% confidence interval). (c) Average* $MSD/(4\Delta t)$. *The dashed line is a fit with Fürth's formula* $MSD = 4D[1 - \frac{P}{t}(1 - \exp(-\frac{t}{P})]$ *with* $D = 11.6 \pm 0.4\ \mu\text{m}^2.\text{min}^{-1}$ *and* $P = 3.0 \pm 0.4$ min. *(d) Distribution of D from individual trajectories, defined as in expression (2.21). The dashed line shows a log-normal fit with* $\langle \ln D \rangle = 1.46$ *and* $\sigma_{\ln D} = 0.90$.

under-estimation of D. Yet, its use is more practical than that of the fit of the MSD or MSD/Δt with any function. Indeed, the latter presupposes either an assumption of the function to fit – most often Fürth's formula – or to choose a fitting range in the case of simple linear fit. To that extent they are not so much more rigorous than definition (2.21), which at least can be used in a systematic way. In particular, even if the 'true' value is not measured, it still allows to compare quantitatively two different values, for instance by taking their ratio. Particular properties, such as the MSD overshoots reported in DH1 [18] or cancer cells, could require more care in doing so, but in the case of AX2 cells, the MSD exhibits a quite standard monotonously increasing behaviour, and even if the plateau is not reached at $\Delta t = 30$ min, the measurements are quite reliable, especially in a relative view.

We checked whether this MSD was well described by Fürth's formula. The result of the fit is shown in Figure 2.3c. Although the line is everywhere 'close' to the data points, it does not give a reliable account of the experimental curve's behaviour: firstly, it underestimates the short-time-scale activity; secondly, at intermediate and large time scales, it converges faster towards a plateau than the actual curve does. This divergence is well-explained by the averaging of two underlying time scales into a single one by Fürth's formula, as we shall see in the next paragraph.

Correlation function. Three avatars of the velocity *direction* autocorrelation function are plotted in Figure (2.4a): the actual VACF, averaged out of all the trajectories; the velocity direction autocorrelation function (VDACF), averaged in the same way; the VACF, first averaged *per experiment*, then over the set of experiments. The last two are strictly indistinguishable, and the inter-experimental noise is an order of magnitude smaller than the intra-experimental one (Fig. 2.4b): this shows that the properties of the correlation function, *ie* the persistence of the cells, is conserved throughout experiments, in spite of the variability in the population and of the variation of other quantities such as the cell speed between experiments.

The normalised VACF takes slightly larger values than the VDACF, but they are still very close to each other. The discrepancies occur from the different weight granted to small or large displacements in the VACF, while VDACF gives all the vectors the same importance. The very limited difference between the two suggests that the speed and persistence dynamics are independent of each other.



(a) *(b)*

Figure 2.4: Velocity autocorrelation function.
(a) Plot of the VACF (green squares) and the VDACF, either averaged out directly from all trajectories (red triangles) or first per experiment, then over the experiments (blue circles). The lines show fits of the VDACF (red triangles): single exponential decay (gray dash-dotted line), double exponential decay (black dashed line). (b) Root-mean-squared deviation of the VDACF. Deviation between the experiments (blue circles), spread inside experiments (red triangles). The error bars show the standard deviation of the latter between experiments.

Both VACF and VDACF exhibit a small drop-out at $\Delta t = \delta t$, and then decrease smoothly to 0 from $C(2\delta t)$. The correlation drop-out at the second point is actually due to the discrete character of the data (see next section), which explains why it is less marked in the VACF (as the small displacements are more sensitive to discreteness).

Trying to fit with a single exponential decay leads to the same issue as the Fürth formula fit of the MSD. Indeed, the exponential curve seems to average out two time-scales: a fast initial decay followed by a slower tail. This is confirmed by fitting (from $\Delta t = 2\delta t$) with a sum of two exponentials, which matches the data much better (Fig. 2.4a):

$$C(\Delta t) = \phi_1 e^{-\frac{\Delta t}{\tau_1}} + \phi_2 e^{-\frac{\Delta t}{\tau_2}} \tag{2.22}$$

Thus our data are a new example of the existence of at least *two* relaxation times in the cells' directionality. Actually, $\phi_1 + \phi_2 < 1$ means that there is an additional, even faster, decay in the correlation, which happens between the first two or three points. We will see that this short relaxation, termed "mode 0", must be considered with care as positional

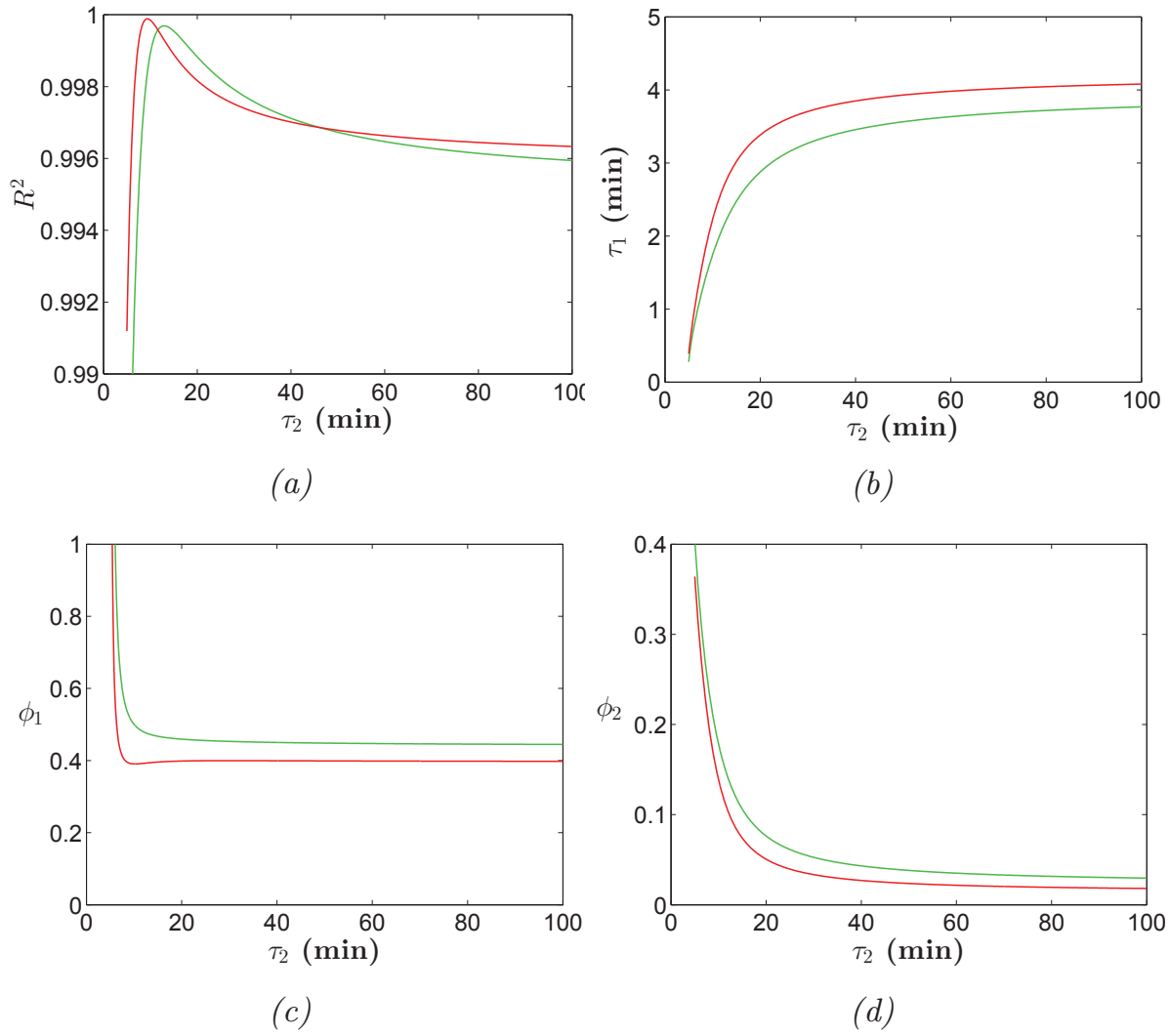


Figure 2.5: Double exponential fitting of the autocorrelation function. (a) R^2-value of the fit of the VACF (green) and the VDACF (red) for fixed τ_2. There is an optimum at $\tau_2 = 12.5$ min for the VACF and $\tau_2 = 9.4$ min for the VDACF. (b) Fitted values of τ_1. (c) Fitted values of ϕ_1. (d) Fitted values of ϕ_2.

noise favours its emergence. Conversely, modes 1 and 2 play a great role in the collective regulation of the motility, as it will appear clearly in **Chapter 3** and more especially in **Chapter 4**.

To study these modes more in depth, we varied τ_2 from 5 to 100 min, while fitting the three other parameters, for both the VACF and the VDACF. They exhibit a maximum in the fit's accuracy at $\tau_2 = 12.5$ min and $\tau_2 = 9.4$ min respectively (Fig. 2.5a). The corresponding values for τ_1 are 2.2 min and 2.1 min respectively (Fig. 2.5b). Later, when it comes about comparing different experiments or different time-point, we use fixed values for τ_1 and τ_2, so that we get comparable values for ϕ_1 and ϕ_2: thus, we gain information about the relative importance

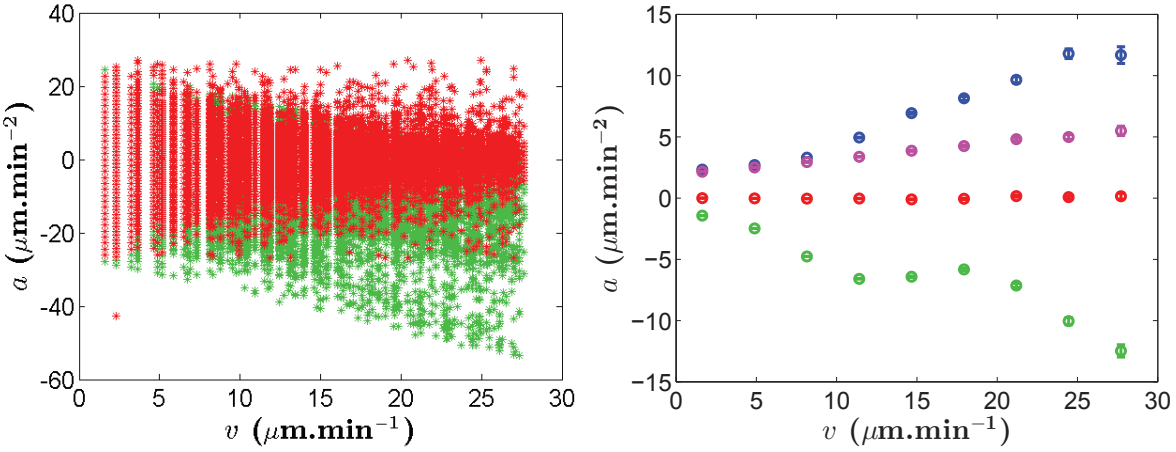
of modes 1 and 2. It seems preferable to miss the precise values of the relaxation times than to waste everything in getting a fuzzy set of parameter values that all vary at the same time. Besides, without a solid underlying model, these precise times would be of no interest.

As soon as it falls under 1, ϕ_1 looks very stable to the variations of τ_2, around 0.45 for the VACF and 0.40 for the VDACF (Fig. 2.5c). ϕ_2 decreases fast with τ_2, because when the persistence is increased, even a small amount of it creates a long tail in the correlation function (Fig. 2.5d).

Acceleration. We analysed the statistics of acceleration in the same way as Selmeczi *et al.* [68]. To that end, for each data point we computed the acceleration in the frame of the velocity vector. All the results are plotted in Figure 2.6a, as a function of the speed v. It is clear that $a_\perp$ fluctuates around 0, while $a_\parallel$ has a decreasing tendency. The mean and root-mean-square deviation (RMSD) of $a_\perp$ and $a_\parallel$ are shown in Figure 2.6b. It confirms that $\langle a_\perp \rangle = 0$, meaning that the cells are not chiral (*ie* they turn with equal probabilities to the left or to the right). Conversely, $\langle a_\parallel \rangle$ decreases linearly with v. Some would see it as an argument in favour of an OU-like process – because it is consistent with the 'drag' term $-\gamma \mathbf{v}$ in the acceleration. On the contrary, in an ABP-like model, one expects that:

$$\mathbf{a} = \dot{v}\mathbf{e}_\parallel + \sqrt{2D_r}\eta\mathbf{e}_\perp \Rightarrow \langle a_\parallel \rangle = \langle \dot{v} \rangle \tag{2.23}$$

In particular, if $v = \langle v_0 \rangle$ is constant, $a_\parallel$ should be 0. Yet, this is a naive view that ignores the discrete aspect of the data. In that case, the measured acceleration is $\mathbf{a}_i = (\mathbf{v}_{i+1} - \mathbf{v}_i)/\delta t$, which has a non-zero parallel component $a_\parallel = v_0(1 - \cos\delta\theta)$ in the frame of $\mathbf{v}_i$, where $\delta\theta$ is the turning angle. Therefore this non-zero $\langle a_\parallel \rangle$ does not rule out an ABP-like behaviour. By contrast, it is clear that the noise amplitude is not constant and anisotropic, showing again that the Ornstein-Uhlenbeck process is not sufficient to account for the whole dynamics of AX2's trajectories.



(a) *(b)*

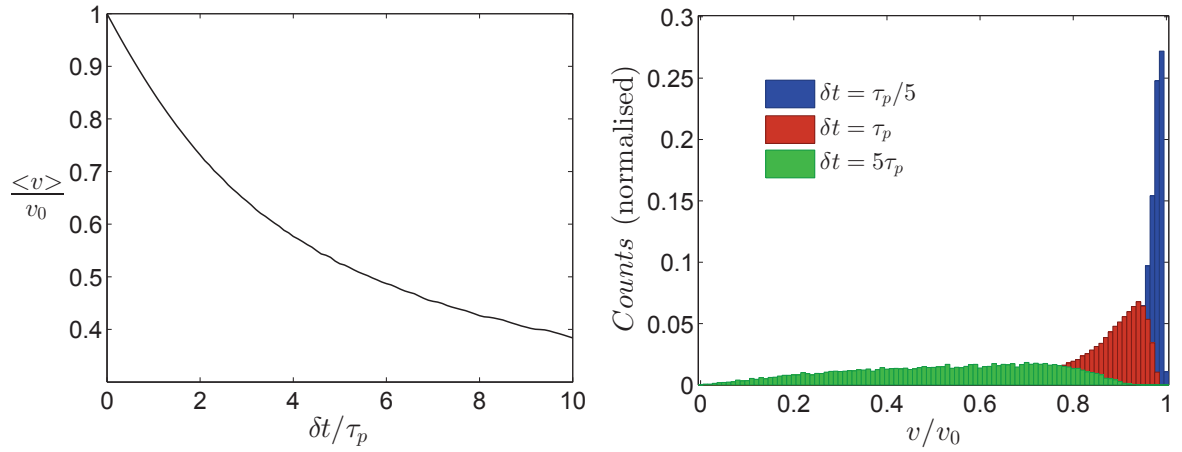
Figure 2.6: Acceleration statistics.
(a) Scatter plot of $a_{\parallel}$ (green) and $a_{\perp}$ (red) as a function of v ($\sim 3 \times 10^6$ points in each scatter). (b) $\langle a_{\parallel} \rangle$ (green), $\langle a_{\perp} \rangle$ (red), and RMSD of $a_{\parallel}$ (blue) and $a_{\perp}$ (purple) as a function of v.

4 Sampling effects and measurement noise: insight from simulations

4.1 Simple active Brownian particles

Contrary to "numerical experiments" in which any precision can be reached, real experiments yield data that have finite time and space resolution. The time resolution is defined by the time lapse. In theory, it could be reduced a lot as we use the camera far from its maximum frequency. Actually, it would not improve the data, due to the limited space resolution. The latter is inherent in the system: it is at least equal to the pixel size in the case of our detection based on the maxima, and even with contour detection the cell's deformations make that the small displacements of the center of mass are not really related to net displacements of the whole cell. We simulated trajectories of active Brownian particles to study the resolution effects on the measurements.

Apparent speed. The problem of the apparent speed have been tackled in some details recently for correlated random walks with exponentially distributed ballistic steps at constant speed v_0 [80]. In this case, during a particular sampling interval, either the particle turn or not. In the second case, the speed is accurately measured by $v = \Delta x/\delta t$, where Δt is the particle's net displacement and δt is the sampling time. On the other hand, if the particle has turned during the interval, $\Delta x < v_0 \delta t$,



(a) *(b)*

Figure 2.7: Apparent speed at various sampling rates.
(a) $\langle v \rangle$ as a function of the sampling time δt. (b) Distribution of the measured v for various sampling intervals: $\delta t = \tau_p/5$ (blue), τ_p (red) and $5\tau_p$ (green).

and so v is underestimated. In the case of continuously turning particles, such as ABPs, this happens during at every sampling interval.

We simulated ABPs moving at $v_0 = 1$ – the space can be rescaled – and with persistence τ_p with small time steps $\tau_p/1000$. Then, the data were discretised by measuring the displacements with varying time interval δt. The plot of the averaged measured speed $\langle v \rangle$ shows that it decreases with the sampling interval (Fig. 2.7a). Yet, when $\delta t < \tau_p$, which usually experimentalists try to achieve, the induced error remains limited, with an approximatively linear decrease towards $v(\delta t = \tau_p) \simeq 0.85 v_0$. The distributions of v for various δt illustrate this in more detail (Fig. 2.7b). Thus, at small δt, the distribution is highly asymmetric and concentrated close to 1. Then it spreads towards 0 with decreasing asymmetry, and when $\delta t > \tau_p$, there is a significant probability to measure any speed between 0 and v_0, the particle's actual speed. However, when $\delta t < \tau_p$, event without a difference of an order of magnitude, the speed measurements are quite reliable (see for instance Fig. 2.7b, blue histogram).

Apparent speed: pixelation effect and positional noise. Now, considering a finite discretisation size δx, a natural speed scale $\delta v = \delta x/\delta t$ appears. It characterises the slowest speed one can detect given the experimental resolution. It results in displacements smaller than δx during δt being either counted as 0, or as δx, hence making an *over*estimation of the speed possible (Fig. 2.8a).

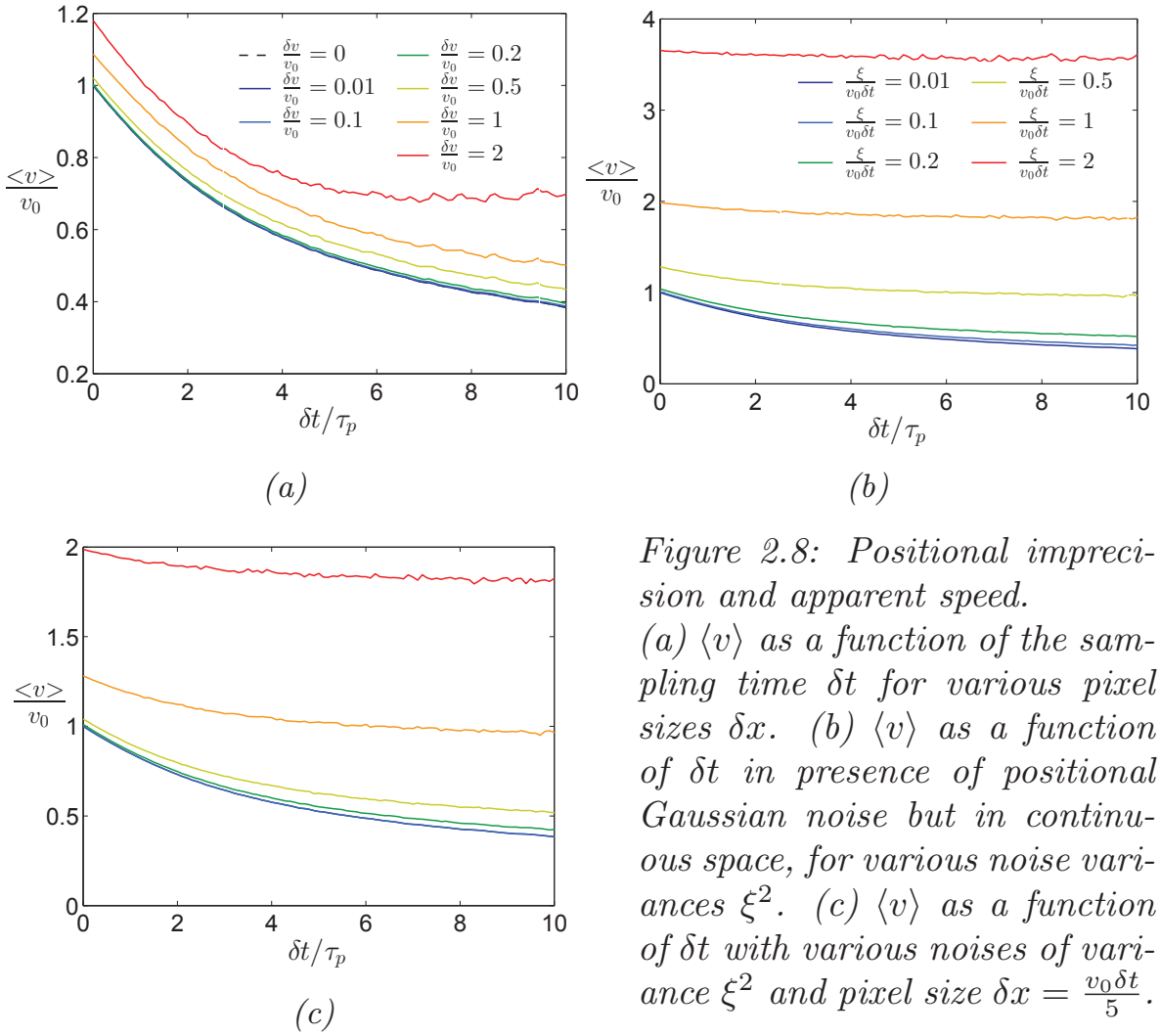
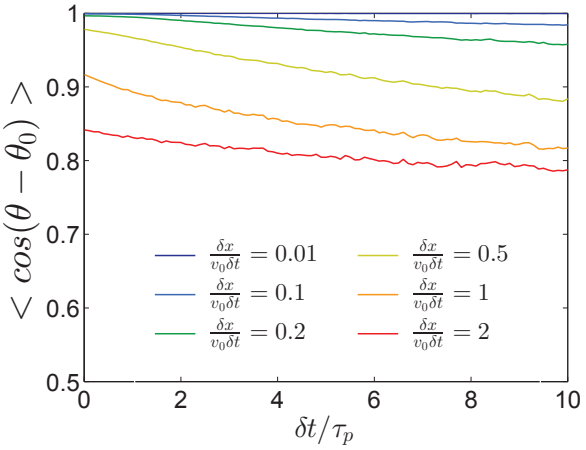


Figure 2.8: Positional imprecision and apparent speed.
(a) $\langle v \rangle$ as a function of the sampling time δt for various pixel sizes δx. (b) $\langle v \rangle$ as a function of δt in presence of positional Gaussian noise but in continuous space, for various noise variances ξ^2. (c) $\langle v \rangle$ as a function of δt with various noises of variance ξ^2 and pixel size $\delta x = \frac{v_0 \delta t}{5}$.

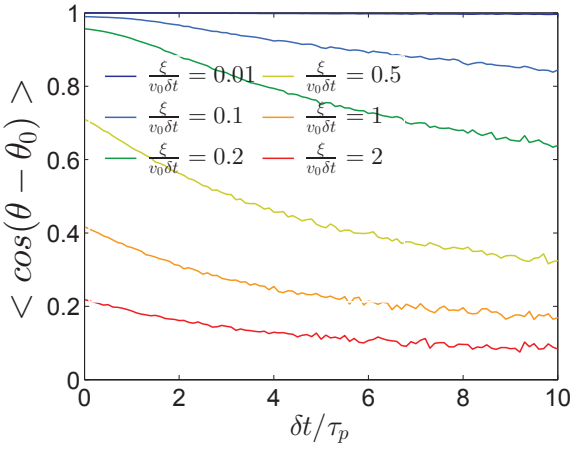
To mimic positional noise, we added an uncorrelated Gaussian noise with variance ξ^2 to the initial continuous-space positions. When ξ approaches $v_0\delta t$, it results in a dramatic increase of the measured v (Fig. 2.8b). Now, taking a pixel size so that $\delta v = v_0/5$ rescues partly this overestimation. Note that in the experiments, typically $\delta x = 1.63\,\mu\text{m}$ and $\delta t = 1\,\text{min}$ **(Chapter 3)** or $\delta x = 0.65\,\mu\text{m}$ and $\delta t = 20\,\text{s}$ **(Chapter 4)**, so that $\delta x/(v_0\delta t) \approx 0.2 - 0.5$. The positional noise ξ is of the order of δx as well (which would not be the case *e.g.* with stage noise).

To conclude, space and time discretisation of the data affects the measurement of v in a way that produces a $\sim 10\%$ error in the experimental conditions. Uncorrelated positional noise would have a much stronger effect. However, it must be stressed out that we only considered *uncorrelated* positional noise. In the case of maximum brightness detection for instance, the fluctuations in detected position are related to the deformations of the cell, which are more likely correlated. As a consequence, to subsequent positions will be in average less distant than if the noise

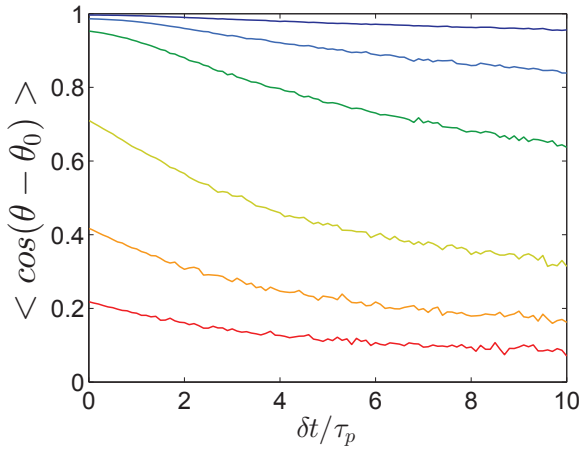
were completely random.



(a)



(b)

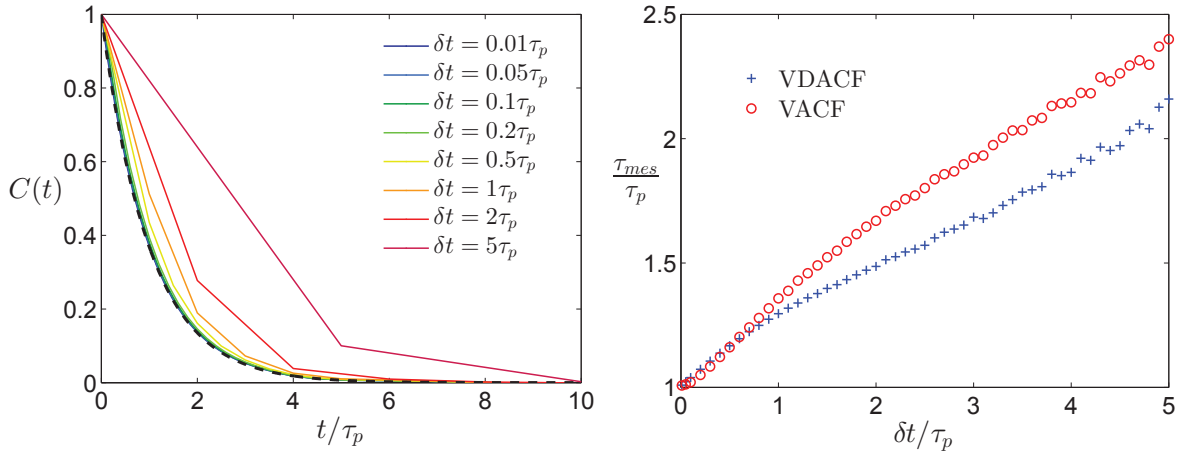


(c)

Figure 2.9: Direction measurement accuracy.
(a) $\langle \cos(\theta - \theta_0) \rangle$ *as a function of the sampling time* δt *for various pixel sizes* δx*. (b)* $\langle \cos(\theta - \theta_0) \rangle$ *as a function of* δt *in presence of positional Gaussian noise but in continuous space, for various noise variances* ξ^2*. (c)* $\langle \cos(\theta - \theta_0) \rangle$ *as a function of* δt *with various noises of variance* ξ^2 *and pixel size* $\delta x = \frac{v_0 \delta t}{5}$*.*

Directional accuracy. Measuring the velocity direction with good accuracy is also fundamental to estimate quantities such as the persistence. We did the same analysis of the effect of spatial imprecisions (resolution and noise) on the measured angle θ with respect to the actual angle θ_0. The plots of $\langle \cos(\theta - \theta_0) \rangle$ (Fig. 2.9) show that, once again, the pixelation yield a reasonable error on the direction, whereas uncorrelated noise has a more drastic effect. This time, though, the accuracy lost because of noise is not rescued by spatial discretisation (Fig. 2.9c). However for $\xi = 0.2 v_0 \delta t$ it is still reasonable, and even a higher error could be overcome by taking an average, as we will see with the correlation function.

Velocity autocorrelation function. We used the same framework to analyse the resolution effects on the correlation function. The time



(a) *(b)*

Figure 2.10: Effect of the sampling time on the correlation function. (a) Autocorrelation function (VDACF) computed for various δt. (b) Persistence time, measured from an exponential fit of the VACF (red circles) or VDACF (blue plus) as a function of sampling time δt.

interval has only a slight effect on the shape of the correlation function, provided $\delta t < \tau_p$ again (Fig. 2.10a). When the correlation function is fitted with an exponential to measure τ_p, the obtained value τ_{mes} increases approximately linearly with δt (Fig. 2.10b), and at $\delta t > \tau_p$ the effect is more marked on the VACF than on the VDACF. At $\delta t = \tau_p$, $\tau_{mes} = 1.38\tau_p$ using the VACF, and $\tau_{mes} = 1.30\tau_p$ using the VDACF, and the overestimation is about 10% for $\delta t = \tau_p/5$, which is close to our typical work conditions.

The pixelation is not too harmful to the correlation function, even if at very large pixel sizes discrete effects start appearing (Fig. 2.11a). The addition of noise is more problematic. As shown in Figure 2.11b, the correlation drops out at $t = \delta t$. The rest of the curve, for $t > \delta t$ is well fitted by an exponential, which it does not tend to 1 any more but to ϕ. We measured systematically τ and ϕ by fitting the autocorrelation function with the expression $C(t) = \phi \exp(-\frac{t}{\tau})$, for pixelised data with various pixel sizes, or noisy data with various noise amplitudes and a fixed pixel size $\delta x = v_0 \delta t/5$. The results are shown in Figures 2.11c–d. While τ is quite well retrieved, ϕ becomes much smaller than 1 when ξ approaches $v_0 \delta t$. This is because this noise adds a fast relaxing mode in the directional persistence. Noisy data without pixelation yield similar results as those presented here (data not shown), so that it is really an effect of noise, and pixelation does not affect the measurements a lot.

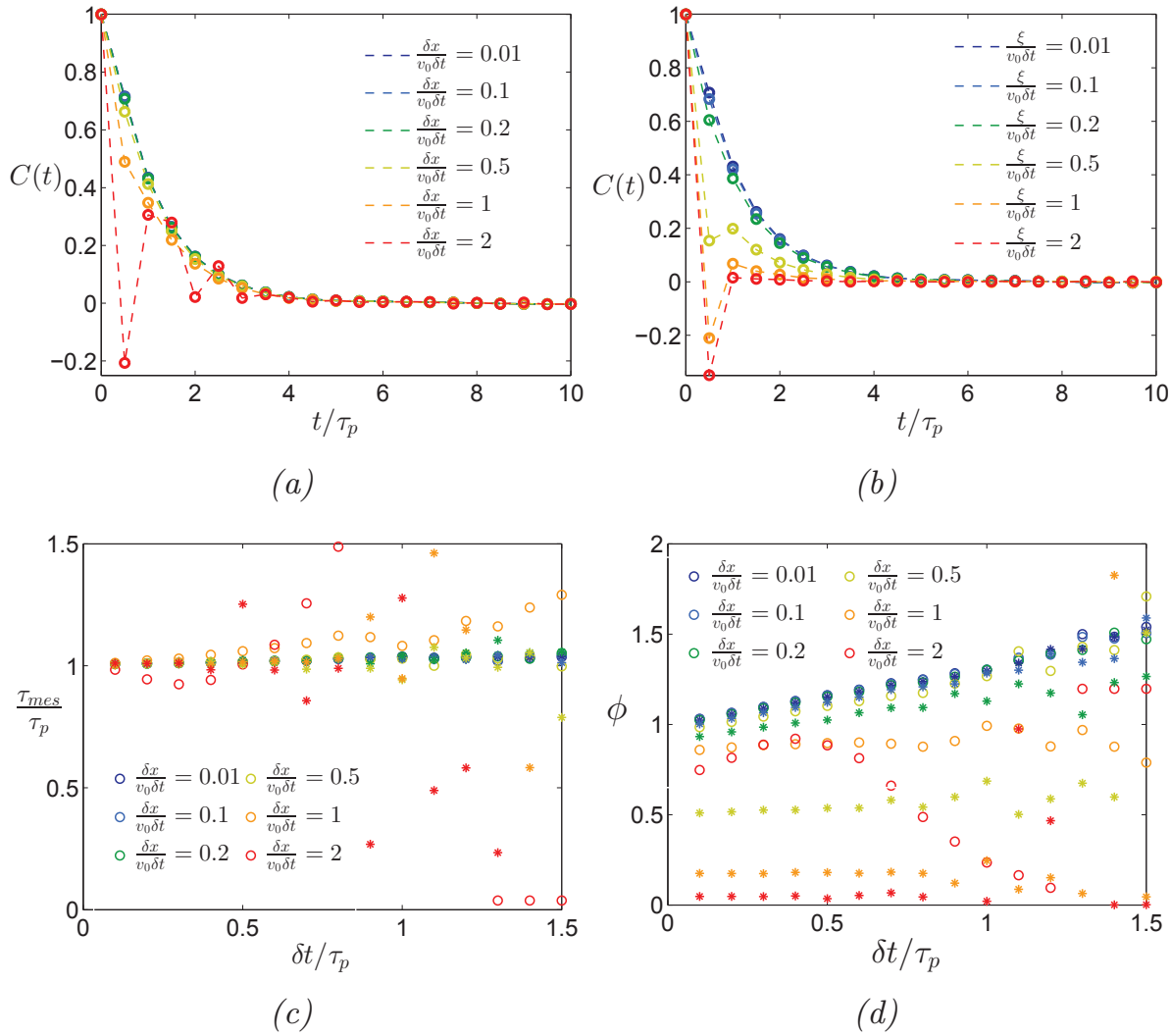


Figure 2.11: Effect of spatial resolution on the correlation function. (a) Correlation function (VDACF) computed for various pixel sizes δx – $\delta t = \tau_p/5$. (b) Correlation function computed for various noise amplitudes ξ – $\delta x = 0.5 v_0 \delta t$, $\delta t = \tau_p/5$. (c) Normalised measured values of τ as a function of δt for various pixel sizes δx (circles) or various noise amplitudes ξ (asters, same colour code as the circles, with $\delta x = 0.2 v_0 \delta t$). (d) Measured values of ϕ, the prefactor of the autocorrelation, as a function of δt for various pixel sizes δx (circles) or various noise amplitudes ξ (asters, with $\delta x = 0.2 v_0 \delta t$).

4.2 Bimodal motion

As we have seen in **Section 3.3**, the velocity autocorrelation function of AX2 cells are best fitted with the function:

$$C(t) = \phi_1 e^{-\frac{t}{\tau_1}} + \phi_2 e^{-\frac{t}{\tau_2}} \tag{2.24}$$

with $\phi_1 + \phi_2 < 1$. Moreover, the first three points exhibit the typi-

cal output of pixelation or positional noise pointed out in the previous subsection. We can suspect from that that there are at least two different time scales in the directional persistence of *Dictyostelium* cells, and maybe shorter third one. There are several means to account for these time scales in a simple model: by integrating correlation in the noise of a monomodal motion (either OU-like or ABP-like), by building an OU-like model with memory kernel *à la* Selmeczi *et al.*, or by postulating bimodal motion. The last option seemed likely to be the closest to the actual cell behaviour (see for instance **Chapter 3 – Section 2.5** or **Chapter 4 – Sections 3.4 and 4.4**). Therefore, we simulated 10^3 bimodal ABPs with the following simplified properties: a "low persistence" mode 1, with persistence time $d_1^{-1} = D_r^{-1} = 2$ min and mean step duration λ_1^{-1}; a ballistic mode 2, with persistence time $d_2^{-1} = +\infty$ and mean run duration $\lambda_2^{-1} = \omega^{-1} = 10$ min. A particle in mode i switches to mode $j \neq i$ with a rate λ_i, resulting in an exponential distribution of step durations. We varied λ_1^{-1} from 1 min to 200 min (*ie* from 0.5 to 100 in units of D_r^{-1}) to cover a wide range of behaviours.

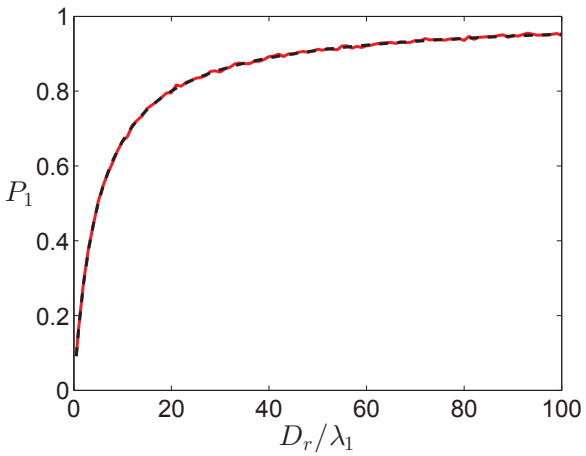


Figure 2.12: Proportion of time spent in mode 1, P_1, as a function of the mean step duration in mode 1 normalised by D_r^{-1}. Measurements from the simulations (solid red line) and theoretical value (black dashed line).

Recovering the input parameters with good time and space resolution. First, we measured the proportion P_1 of the trajectories spent in mode 1. It compares well to the theoretical value $P_1 = \frac{\omega}{\omega+\lambda_1}$ (Fig. 2.12).

Then we computed the velocity autocorrelation function for each of the simulation and extracted the parameters from a fit $C(t) = \phi_1 \exp(-\frac{t}{\tau_1}) + \phi_2 \exp(-\frac{t}{\tau_2})$. First, we allowed the four parameters to take any value without any initial guess (Fig. 2.13a-c). Then, we set the starting point at 0.5 for both ϕ_i, 2 min for τ_1 and 20 min for τ_2 (Fig. 2.13b-d). We compared the extracted values to their theoretical predictions (see **Appendix A** for the detailed expressions). They are quite well recovered,

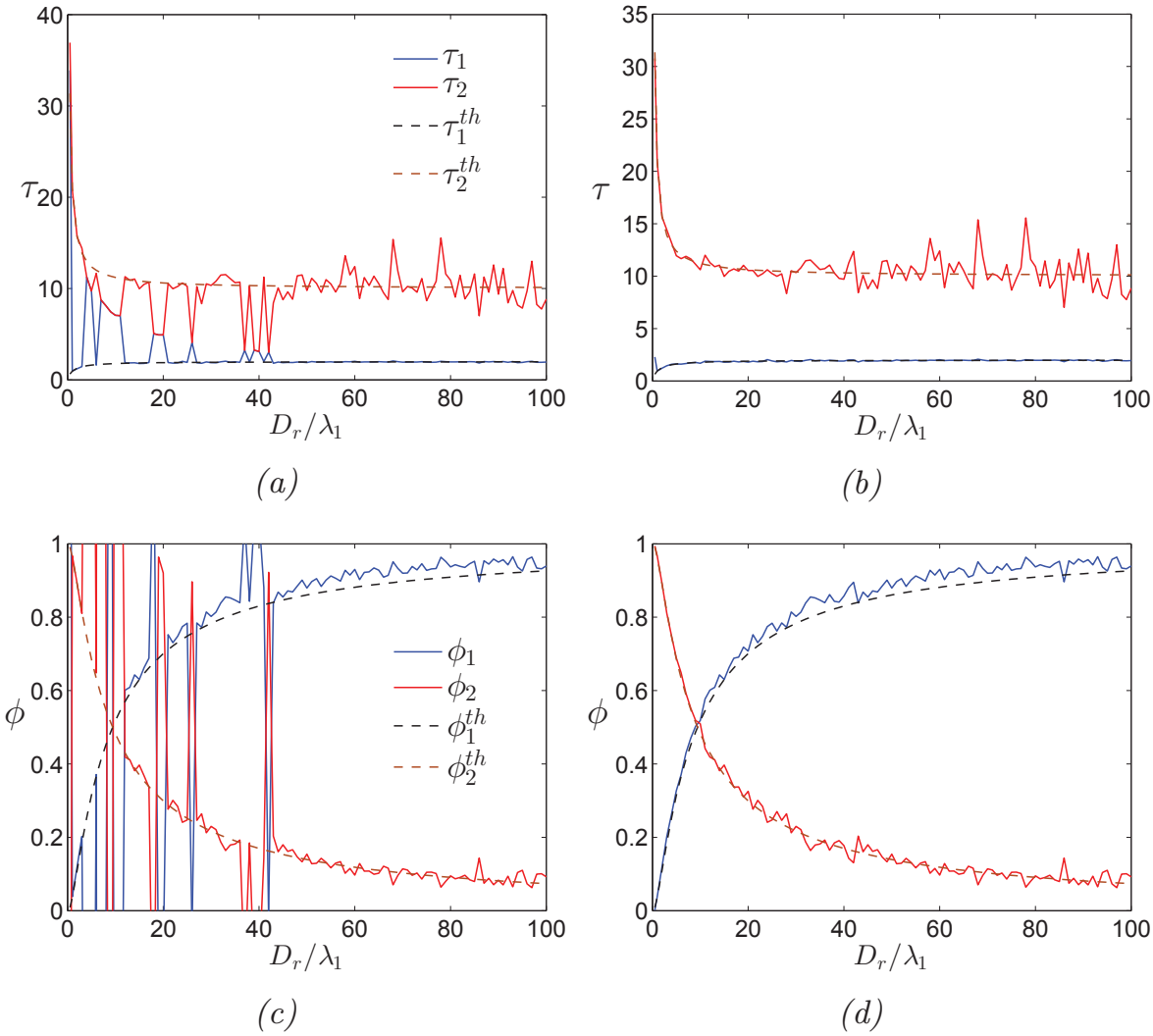


Figure 2.13: Parameters of the autocorrelation function.
(a) τ_1 (blue) and τ_2 (red) measured from the autocorrelation functions for various λ_1^{-1} (solid lines), and comparison to theoretical predictions (dashed lines). (b) τ_1 and τ_2 measured from the autocorrelation functions with initial guesses (solid lines) and comparison to theoretical values (dashed lines). (c) ϕ_1 and ϕ_2, measurements from the free fit (solid) and theoretical (dashed). (d) ϕ_1 and ϕ_2, measurements from the fit with initial guess (solid) and theoretical (dashed).

but with increased reliability in the second case.

Because of the various sources of error, it can be a risky to let all the parameters be set freely by the fitting procedure in the case of experimental data. Although τ_1 and τ_2 are not strictly equal to 2 min and 10 min respectively, we chose to test the results of the two-parameters fit $C(t) = \phi_1 \exp(-\frac{t}{2}) + \phi_2 \exp(-\frac{t}{10})$. The results are in quite good agreement with the expected values for $\lambda_1^{-1} > D_r^{-1}$ (Fig. 2.14). This is not surprising as $\tau_1 \simeq 2$ min and $\tau_2 \simeq 10$ min in this range. Conversely, the

measured ϕ_i diverge quite fast from the true values when $\lambda_1^{-1} < D_r^{-1}$.

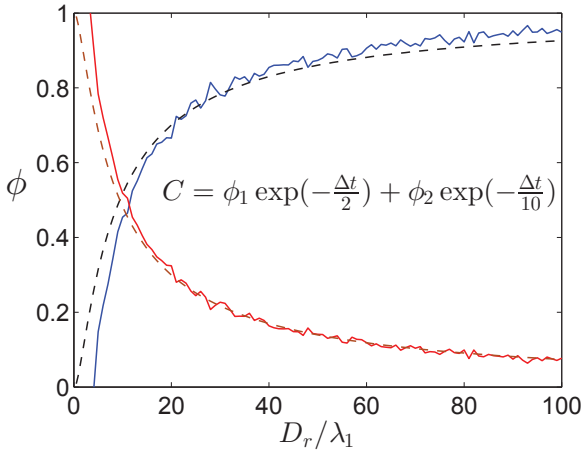


Figure 2.14: ϕ_1 (red) and ϕ_2 (blue) extracted from a fit where $\tau_1 = 2$ min and $\tau_2 = 10$ min are constrained (solid lines), and comparison to theoretical predictions (dashed lines).

Resolution effects on the fit results. We did the same analysis after discretising the data as in **Section 3.4**. In particular, as shown in Figure 2.15, a reasonable sampling up to $\delta t = D_r^{-1}/2$ does not affect much the results. This is also the case for pixelation (Fig. 2.15a-c). By contrast, again, positional noise introduces a faster decorrelation, hence a decrease of the measured ϕ_i values (Fig. 2.15b-d). Importantly, for a particular noise amplitude, the ratios ϕ_1/ϕ_1^{th} and ϕ_2/ϕ_2^{th} are equal and constant. It means for a given source of noise, its influence in the correlation function *does not depend* on the details of the trajectory. In other words, when a "mode 0" is detected, its amplitude $\phi_0 = 1-(\phi_1+\phi_2)$ should remain constant. If not, it could mean that the "mode 0", or a part of it, is intrinsic to the trajectory.

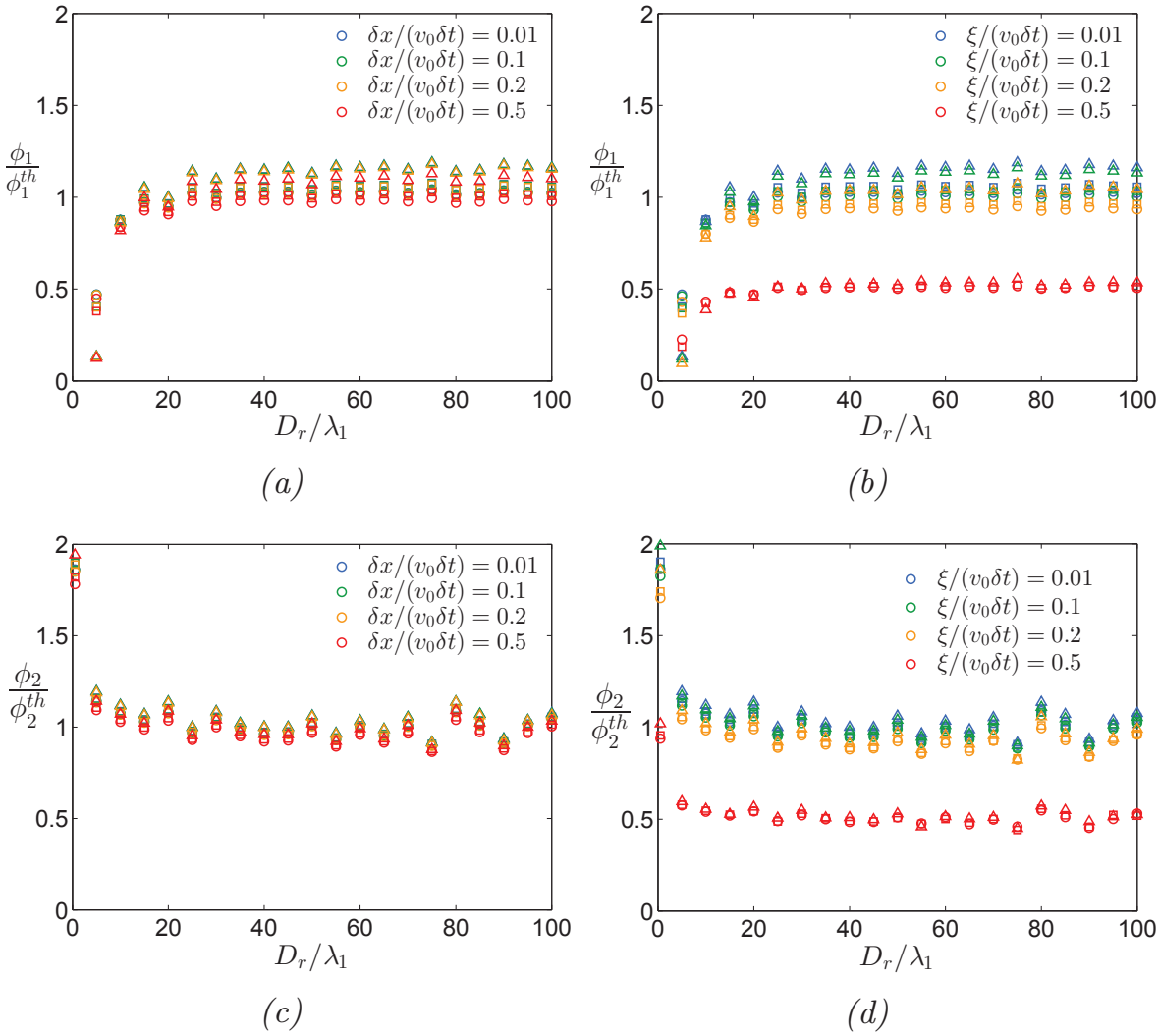


Figure 2.15: Sensitivity of ϕ to spatial resolution.
(a) ϕ_1/ϕ_1^{th} for $\delta t = 0.1D_r^{-1}$ (circles), $0.2D_r^{-1}$ (squares) and $0.5D_r^{-1}$ (triangles) and various pixel sizes δx. (b) ϕ_1/ϕ_1^{th} for $\delta t = 0.1D_r^{-1}$ (circles), $0.2D_r^{-1}$ (squares) and $0.5D_r^{-1}$ (triangles) and various noise amplitudes ξ. (c) ϕ_2/ϕ_2^{th} for various δt and pixel sizes δx. (d) ϕ_2/ϕ_2^{th} for various δt and noise amplitudes ξ.

5 Conclusions

Cell trajectories are a quite simple observable that can be obtained without the need for sophisticated imaging. They characterise the cells' motion, and result from the interplay of intracellular noisy dynamics and environmental cues. As a result, they exhibit the properties of correlated random walks, in particular in the long time limit. At short time scales, however, they fall outside the scope of the classical simple models.

In particular, we have demonstrated in this thesis that the velocity autocorrelation function of our *Dictyostelium* cells exhibit not one but two different relaxation times. To be able to interpret the associated measurements in a reasoned way, we calibrated them on the simplest model that includes smooth direction changes, namely the ABP model. We quantified the effects or discretisation of the data and of random positional error for monomodal and bimodal ABP trajectories. We found that despite an error introduced by these biases, the speed is quite accurately estimated in the experimental conditions. By contrast, the velocity autocorrelation function, or rather the VDACF which we focus on in the experimental data analysis, is strongly affected by the spatial biases: it is only usable from the point at $2\delta t$, and introduces an apparent "mode 0" that decorrelates faster than δt. Yet, the amplitude of this mode 0 remains essentially constant as long as its source does: it means that the measured relative amplitudes of modes 1 and 2 (the real, more slowly relaxing modes) are reliable, and even that changes in the observed amplitude of mode 0 could be ascribed to the inherent dynamics of the cell itself.

We also tackled the problem of trajectory segmentation by bimodal analysis (see **Appendix B**), looking for a scheme that would allow distinguishing between high and low persistence portions of trajectories. We described such a process, but found that the punctual mode allocation errors that occur from it could destroy the distribution of step durations, hence complicating any further analysis and making its results unreliable as such, at least to trace back the underlying statistics of step durations. A method based on Bayesian inference, recently proposed by Metzner *et al.* [64] and that mention also in **Appendix B**, seems promising in the view of measuring an "instantaneous" persistence time, but it also needs more development.

Therefore, in the current state of our knowledge, we can measure the cells' speed and diffusion constant with good accuracy. We can estimate the relative amounts of persistence modes in the trajectories by extracting ϕ_1 and ϕ_2 from the VDACF, using a two-exponential fit with

fixed time-constants τ_1 and τ_2 (at 2 and 10 min respectively). The CME computed at various time lags Δt allows us to estimate the persistence without integrating long portions of many trajectories: it can be used at a more instantaneous and local scale, but it does not relate rigorously and quantitatively to the underlying persistence time. Last, as far as we can judge, bimodal analysis is still too biased to provide additional reliable information on the statistics of cell trajectories (see **Appendix B**). In **Chapters 3** and **4**, we use this framework extensively to characterise the collective regulation of the motility in *Dictyostelium discoideum* cells.

---

3. Regulation of the cell motility by a secreted "Quorum Sensing Factor"

In this chapter, we report on the experimental study of a "Quorum Sensing Factor" (QSF) that is secreted by *D.d.* cells and regulates their motility. The existence of such a factor was demonstrated previously in Laurent Golé's PhD thesis for DH1 cells. Here, we extended the study to the AX2 strain, which we focus more on in the present work, and started deciphering the nature and mode of action of the QSF. In particular, we showed that it is a secreted amphiphilic molecule of molecular weight $\sim$ 10 kDa, which is not charged in experimental conditions and with no dominant proteinic part. We also studied the non-linear response of the cells to the QSF concentration, and could thus indirectly measure its secretion dynamics. We found that it is negatively regulated with saturation, which could be a surprisingly robust way for the cells to compute the cell density. Then, we explored some possible signalling pathways for the response to the QSF, demonstrating the involvement of receptor-bound complexes but not of the cAMP pathway. Lastly, we looked in more detail at the effect of QSF on the cell trajectories' properties and on the single cell dynamics.

Contents

1 Introduction

In **Section 4** of the **introduction part**, we saw how important the ability to sense the others at a distance is, for individuals in a group. The quorum-sensing (QS) systems provide such ability to cell populations. First described in bacteria, they consist in molecules, termed "Quorum-Sensing Factors" (QSF) that are both secreted in the environment and detected by the cells [81]. The concentration of such QSF thus increases with the cell number or density, allowing the group to react accordingly. Most of the known QSF act to regulate the cell proliferation in bacteria, fungi [82] and amœba [83], but also in higher eukaryotes. Such mechanisms could be at stake *e.g.* in tumour dormancy [20] or fibrosis diseases [11].

Quorum-sensing in *Dictyostelium discoideum*. Several QS systems have been observed in *D.d.*, which provides a good model for this, thanks to its both rich and quite simple life-cycle. When they grow on bacteria, *D.d.* cells secrete a glycoprotein called prestarvation factor (PSF), at a constant rate [83]. The comparison of PSF concentration to bacteria number leads to a prestarvation behaviour when the food supply becomes insufficient [84]: the proliferation stops and early developmental genes that prepare the cells for aggregation start being expressed [85, 86]. Importantly, PSF has never been observed in axenic growth conditions. Upon starvation, the cells secrete another glycoprotein named Conditioned Medium Factor (CMF), which, together with PSF, triggers aggregation by potentiating cAMP emission and detection [87, 88]. During late aggregation, counting factors (CF) control the aggregate size through the regulation of motility and adhesion [89].

Another QS system for *D.d.* in nutrient-rich conditions has been studied in the last decade [9, 11]. It is constituted of two main secreted proteins, AprA and CfaD that have two effects. First, they inhibit the growth by lowering both the proliferation rate and the carrying capacity [11]. Then, AprA acts as an autocrine chemorepellant, for which activity CfaD is also needed [22].

During his PhD, Laurent Golé discovered a novel type of QSF in vegetative DH1 cells [18]. He showed that such a factor, secreted during the growth in HL5 medium, was responsible for a substantial density-dependent decrease of the cell motility. This effect is unambiguously detectable from the average diffusion constant taken over at least hundred cells, although the individual values are widely distributed. He gathered data from many experiments of different kinds:

- "medium aging": evolution of the motility over time, while the QSF is continuously secreted;
- "conditioned medium": measurement of cell motility in a medium where cultured cells were previously allowed to grow and secrete QSF;
- "medium perfusion": the QSF was rinsed out by a continuous flow of fresh HL5 medium (FM).

Assuming a constant secretion rate α and rapid homogenisation of the medium, he was able to compute the QSF concentration c up to a factor α with a very simple kinetics model. The plot of the diffusion constant D *versus* the concentration c shows a quite acceptable collapse of all data onto an empirical master curve (Fig. 3.1) whose equation writes:

$$D(c) = D_0 \left(1 + m \exp(-\frac{c}{c_0})\right). \tag{3.1}$$

Here, D_0 is the cells' diffusion constant in pure HCM, mD_0 is the diffusion constant in pure FM and c_0 is a sensitivity threshold for the QSF concentration.

To our knowledge, such a regulation of the motility through a secreted QSF had not been reported before. We decided to pursue the study in order to better understand the chemical nature and the mode of action of this unknown QSF. Before we present our results, let us introduce the very simple and systematic approach that we adopted. All the experiments in this chapter are motility assays that follow the same protocol. We prepared home-made plastic wells stuck on a 1 mm-thick, ethanol-washed glass slide (90°ground edges, precleaned, plain, Marienfeld GmbH, Germany). After UV sterilisation of the sample, 800 µL of cell suspension at 5.7×10^3 mL^{-1} were deposited in each well, to reach a surface density of $\sim 2 \times 10^3$ cm^{-2}. The cells were allowed to sediment and adhere for 45 min in the incubator before medium change and imaging for at least 100 min by low-magnification defocused microscopy (3X or 4X magnification depending on the experiments, and $\delta t \simeq 1$ min). The principle of most of the experiments, with the notable exception of those of section 5 of the present chapter, was to measure the average diffusion constant D in various test conditions, and to compare it to the control values in both fresh and highly conditioned medium (FM and HCM respectively).

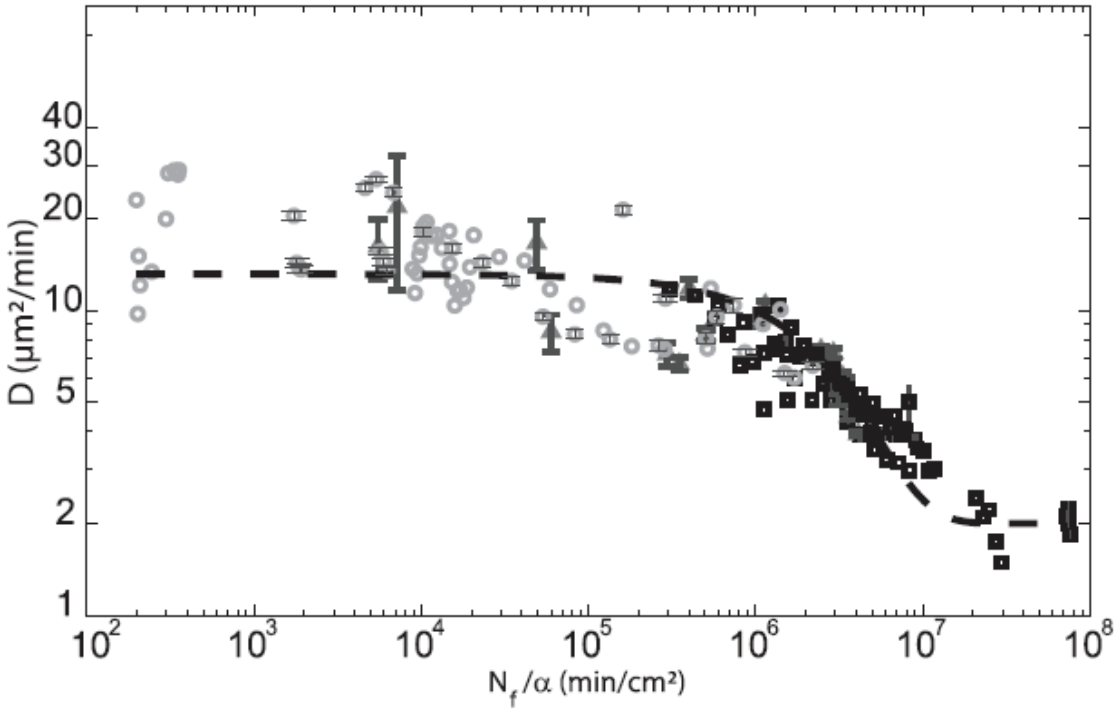


Figure 3.1: Dictyostelium cells secrete an unknown chemical factor that regulates its motility (from [18]).
An empirical relation was found between the calculated concentration N_f/α of this putative QSF and the diffusion constant D of the cells.

Depending on the context, three different quantities that all measure the amount of QSF activity may be used. The most obvious is the raw value of D. Yet the relative diffusion constant defined as $R = \frac{D}{D_{ref}}$, where the reference taken is most often the value of D in FM, denoted D_{FM}, has some advantages over it: first, it spares the display of D_{FM}, which in most cases does not present much interest by itself; then it allows an easy comparison of situations with different basal motility, namely D_{FM}. This can happen *e.g.* when different strains are compared, due to the reference medium, or just by chance. Indeed, we observed the general trend that a poorer medium leads to a general increase of the motility, as a smooth transition between nutrient-rich and starved conditions. Furthermore, we sometimes observed surprising changes of D_{FM} between experiments, usually spaced in time but also rarely inside a single series of experiments. This effect, although known from the *Dictyostelium* community, remains unexplained. However, the QSF response is robust to it, so we will not elaborate much on this topic. A typical value for the relative diffusion constant R in HCM is in the 0.3 – 0.4 range. Last, the QSF activity A, defined as

$$A = \frac{D_{FM} - D}{D_{FM} - D_{HCM}} \tag{3.2}$$

allows a rigorous comparison of intermediate D values from different experiments, provided the reference FM (fresh medium) and HCM (highly conditioned medium) used are the same. It varies from 0 in FM to 1 in HCM, with possible negative values due to fluctuations around D_{FM} in absence of QSF, and $A > 1$ if the reference HCM is not conditioned enough. It is mostly used in section 3, to quantify the cellular response to varying concentrations of QSF.

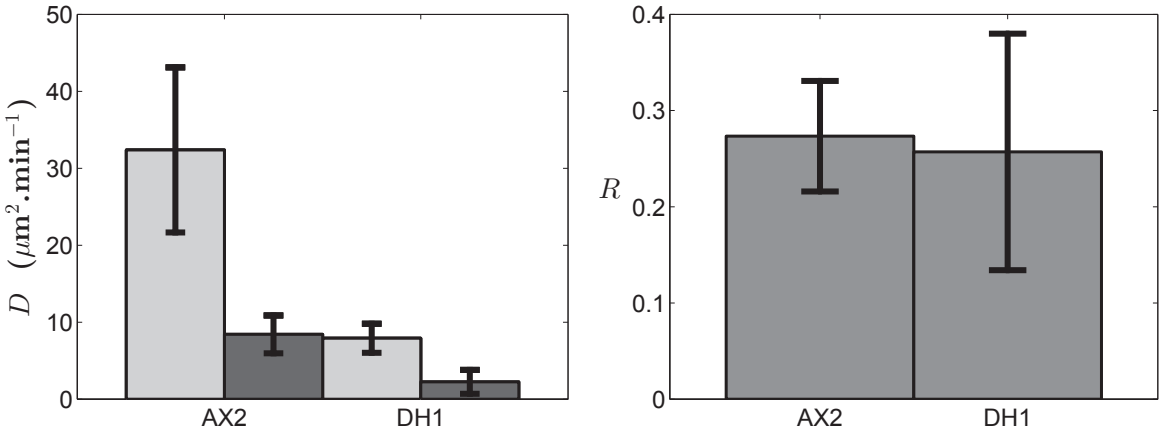
Preliminary experiments. First, as this thesis' main model is the AX2 strain, we extended L. Golé's main result to these cells: using HCM made from both DH1 and AX2 cells, we showed that the QSF effect is shared by the two strains – as it is also with the JH8 strain (see section 4). Interestingly, the relative diffusion constant seems independent from the basal strain's motility: thus, in this experiment, AX2 cells were more motile than DH1, but the ratio R is the same for both strains. More importantly, the QSF seems not to be strain specific: we did not test compatibility of HCM from AX2 on DH1 cells and reciprocally, but as shown in section 4, HCM prepared with an AX2 culture exhibits a normal QSF effect on JH8 cells, which are very similar to DH1.

After these preliminary experiments, we looked in more detail into the properties of the QSF. Firstly, we characterised the biochemical properties of the QSF. Secondly, we analysed the response function of the cells and their secretion dynamics. Thirdly, we explored some possible signalling pathways of the cellular response to the QSF. Lastly, we studied the effect of the QSF on the cell trajectories and on the cell shape dynamics to better understand its mode of action.

2 Biochemical characterisation of the QSF

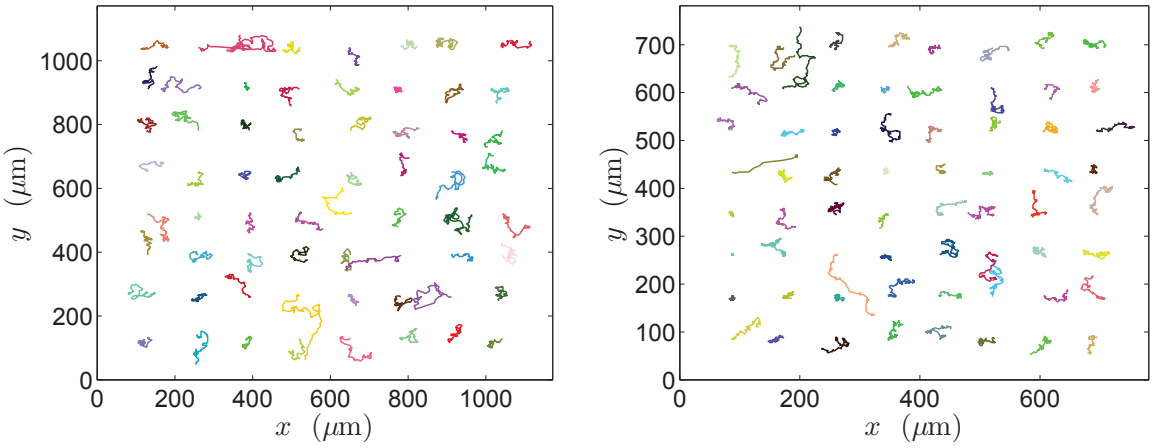
2.1 Estimation of the molecular weight by ultrafiltration

To start with a rough characterisation of the QSF, we conducted ultrafiltration experiments. In these, the media were filtrated through nanoporous membranes (Vivaspin, Sartorius Gmbh, Germany) by means



(a) (b)

Figure 3.2: Comparison of the motility and the QSF response of AX2 and DH1 cell lines.
(a) D for AX2 (left) and DH1 (right) cells in FM (light gray) and HCM (dark gray). AX2 cells appear more motile than DH1 cells. (b) Relative diffusion constant R in both strains: despite their different motilities, the two strains exhibit the same ratio, hence they respond similarly to the QSF.



(a) (b)

Figure 3.3: Typical trajectories of AX2 cells.
Trajectories of AX2 cells in (a) FM and (b) HCM. They do not look qualitatively very different, but they are more contracted in HCM (NB: the scales are different). The origins of the trajectories have been put on a regular array for the sake of clarity.

of centrifugation at 6000g, resulting in the separation of a filtrate fraction containing only the molecules smaller than the membrane's molecular weight cut-off (MWCO) and a concentrate fraction, containing all the

initial constituents but with those of molecular weight higher than the MWCO concentrated up to a factor 10. Hereafter, we will use subscripts following the medium name to denote these fractions. For instance, HCM_c:FM_f refers to a mixture of concentrate of HCM and filtrate of FM. Unless otherwise specified, the concentration factor is 10, and so one volume of concentrate is mixed in 9 volumes of filtrate to recover the initial concentrations of species (Fig. 3.4).

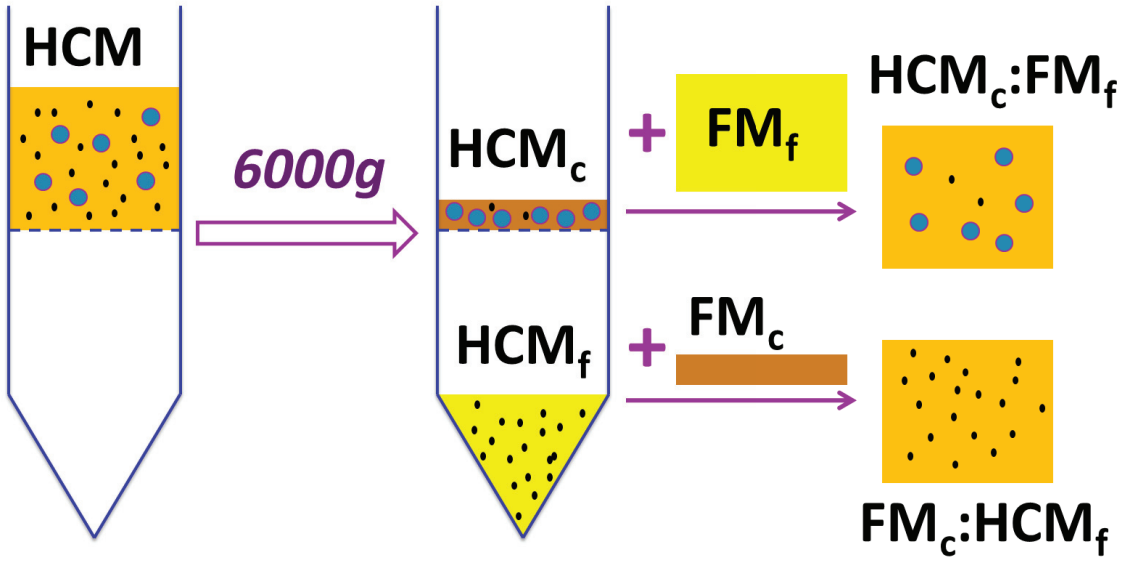


Figure 3.4: Principle of the ultrafiltration and blends.
Starting from HCM where large blue and small black molecules have been secreted, we arrive at HCM_c with concentrated blue molecules and the same concentration of black molecules, and HCM_f with no blue molecule and the same concentration of black molecules as initially. When these fractions are blended with the complementary fraction from FM, one gets HCM_c:FM_f with conserved blue activity but decreased black activity; and FM_c:HCM_f with conserved black activity but lost blue activity.

The QSF is a secreted factor of large molecular weight. First, using a 3 kDa MWCO, we confirmed that the QSF is actually secreted and that its main constituent is larger than 3 kDa. As shown in Figure 3.5, the ultrafiltration procedure does not bias the measurement. We made "reconstituted" rFM and rHCM, respectively FM_c:FM_f and HCM_c:HCM_f. These are just a reassembly of both separated phases of the same medium, and they yield results that are very similar to those of the original media. Only in rFM is the motility slightly increased with respect to FM. This is likely because some molecules are trapped on the ultrafiltration membrane. Yet this does not affect the results greatly, and in particular the response to the QSF is conserved.

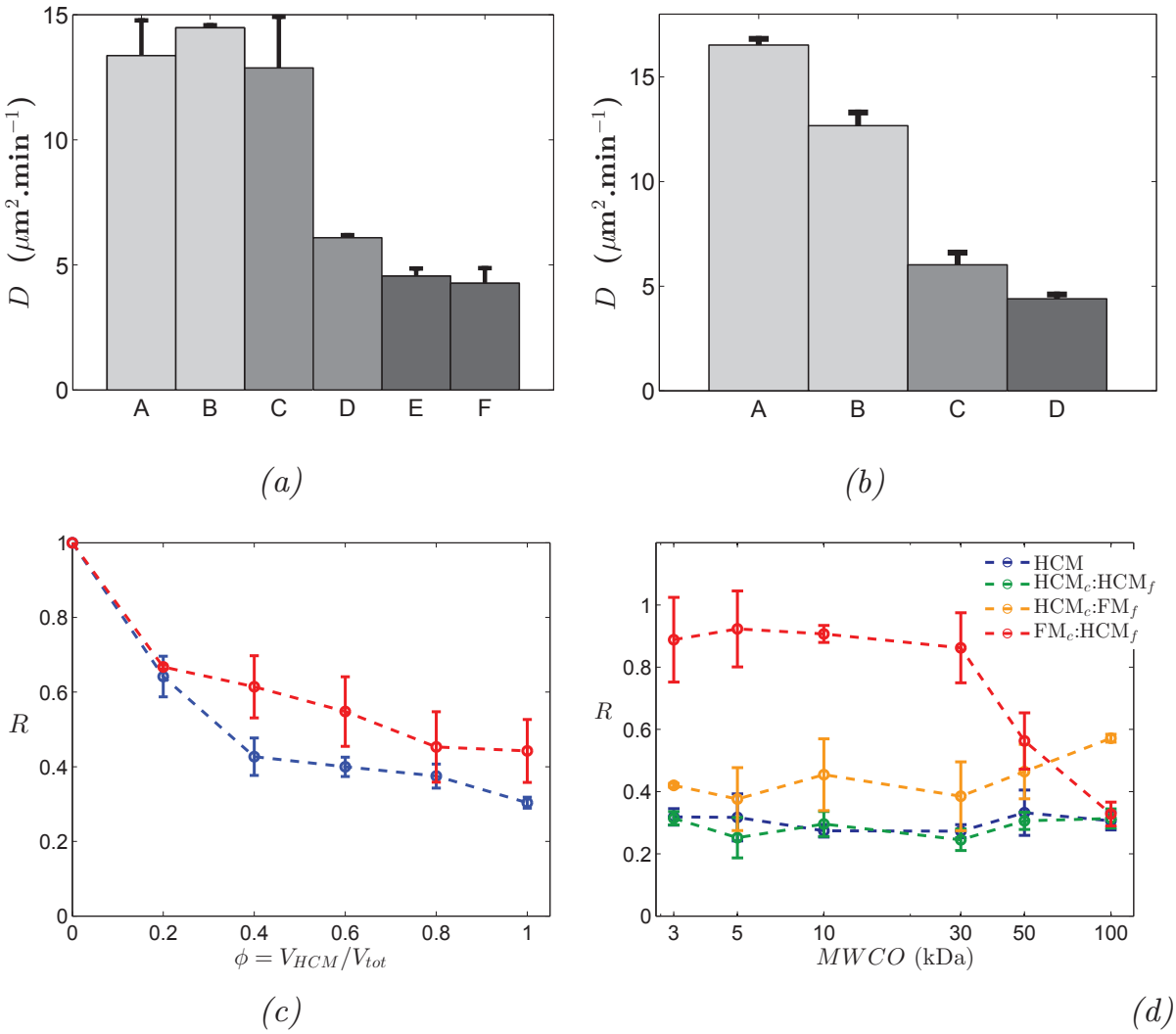


Figure 3.5: Ultrafiltration results.
(a) D in various media, for a 3 k*Da MWCO. A: FM, B: rFM, C: FM_c:HCM_f, D:HCM_c:FM_f, E: rHCM, F: HCM. (b) D in various media, for a* 3 k*Da MWCO. A: FM, B: FM_c:FM, C: HCM_c:FM, D: HCM. (c) Relative diffusion constant R as a function of QSF dilution, either by diluting HCM in FM (blue) or HCM_c in FM_c (mixed at 10% in FM_f, red), for a* 3 k*Da MWCO (d) QSF relative diffusion constant as a function of the cut-off size (HCM_c:HCM_f=rFM). There is a clear cross-over at* $MWCO \simeq 50\ 10^3 Da$.

Now, using FM_c:HCM_f and HCM_c:FM_f we found that most of the QSF activity is contained in the HCM_c fraction. This means that the QSF is actually a large secreted molecule. Indeed, there are four different options: it can be (*i*) a small depleted molecule, (*ii*) a small secreted molecule, (*iii*) a large depleted molecule, or (*iv*) a large secreted molecule. Cases (*i*) and (*ii*) are ruled out by the experiment in Figure 3.5a: a small depleted molecule would be rescued in condition D by the presence of FM_f, whereas a small secreted molecule would be absent in

D but present in C, leading to converse results. Case (*iii*) is ruled out by the experiment with full FM (Fig. 3.5b), as a large depleted molecule would be rescued in condition C by the presence of a large amount of FM (and likely also in condition C of Fig. 3.5). Then, case (*iv*) – large secreted molecule – is compatible with all the results: the QSF present in the HCM is suppressed from the HCM_f by the filtration, but it is concentrated in HCM_c (Fig. 3.5a, bars C–D and 3.5b, bar C).

We ran a dilution experiment (Fig. 3.5c), which confirms these results. Here, we classically diluted HCM in FM on one hand. On the other hand, we diluted HCM_c in FM_c so that the HCM_c:FM_c ratios correspond to the HCM:FM ratio. Then the HCM_c:FM_c mix was diluted in FM_f in a 1:9 ratio, so that the final concentrations of QSF in the simple HCM:FM mixture and in the HCM_c:FM_c:FM_f mixture was the same. At a given QSF dilution ϕ, both yield a similar relative diffusion constant showing that both methods lead to the same QSF concentrations, and again that QSF is completely contained in the HCM_c fraction.

Notice that the QSF response is not completely recovered in HCM_c:FM_f, with $A = 0.77$ (Fig. 3.5a, bar D and 3.5c). That may mean that a factor of small molecular weight is needed to completely activate the main QSF part with $M > 3$ kDa.

Estimation of the size of the QSF. To estimate the size of the QSF, we re-did the ultrafiltration experiment with varying MWCO. We found (3.5c) that around 50 kDa, the QSF activity switches from HCM_c to HCM_f. This means that $M_{QSF} \sim 50 - 100$ kDa. The largest MWCO for which full response is kept in HCM_c is 30 kDa.

Note that we have used the molecular weight as the measurement of the QSF size, as it is the information provided by the manufacturer. Yet, the real parameter of the membrane is their pore radius, which should so correspond to the Stokes radius of the filtrated species. The $MWCO$ are given for *proteins*, for which it relates with the Stokes radius as $R_{glob} \sim M_{glob}^{\frac{1}{3}}$. For linear polymers in contrast, the relation is $R_{lin} \sim M_{lin}^{\frac{1}{2}}$. Thus, if the QSF was a linear polymer and not a globular protein, its actual molecular weight would be $M_{lin} = f.MWCO^{\frac{2}{3}}$, where f is a numerical factor computed from the graph in Figure 3.6, and in that case, $M_{QSF} \sim 10 - 30$ kDa.

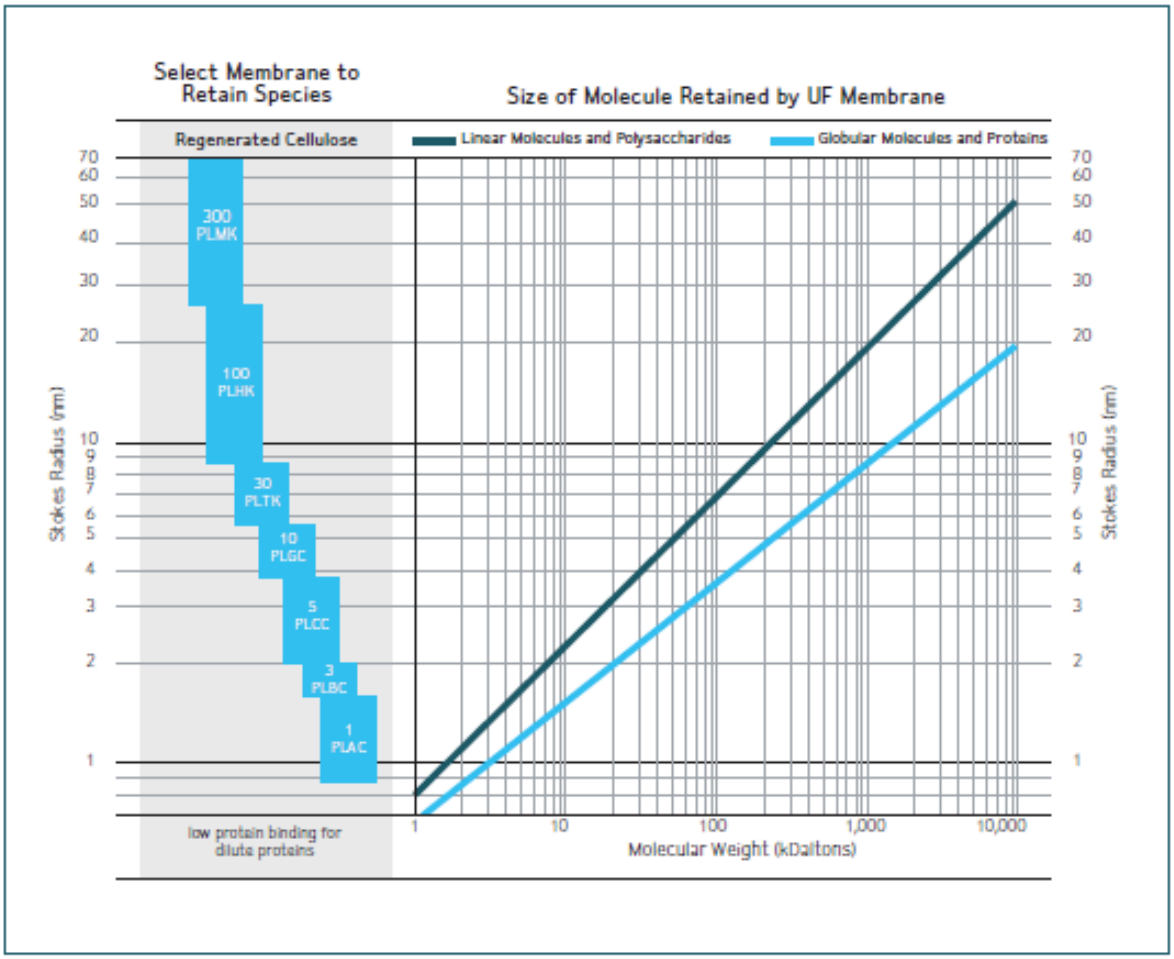


Figure 3.6: Correspondance of MWCO to Stoke's radius and linear polymer molecular weight (from www.millipore.com). A MWCO of 50 k*Da corresponds roughly to a molecular weight* $M_{lin} = 10$ k*Da for a linear polymer.*

2.2 Biochemical characterisation

To get closer to identify the QSF, we ran a series of tests to uncover its chemical properties. The molecular weight estimated in the previous subsection is characteristic of proteins, but some carbohydrate polymers can also reach such – quite big – size. Other candidates could be glycoproteins or proteoglycanes, which are combinations of proteins and sugars, non-covalent aggregates of smaller molecules, or even, although less likely, lipids or nucleic acids.

The QSF is depleted by phenol-chloroform extraction. HCM and FM were subjected to the following treatment:

(i) 20X concentration through a 30 kDa MWCO;

(ii) extraction by a phenol-chloroform mixture;

(iii) two additional extractions with chloroform only to eliminate the phenol dissolved in the aqueous phase;

(iv) 10X dilution in ultra-pure water;

(v) 20X concentration through a 3 kDa MWCO;

(vi) 40X dilution of the remaining in non-treated FM.

The phenol-chloroform extraction is a classical liquid-liquid extraction technique to purify nucleic acid samples. Indeed, the phenol acts by denaturing the spatial structure of the polymers in the solution, then the amphiphilic ones migrate to the organic-aqueous interface where they form a precipitate. The extracted molecules include almost all the proteins, but also some carbohydrate polymers and some lipids. DNA is not extracted in neutral or slightly alkaline conditions ($pH \sim 7-8$), but it is in acidic conditions such as in our experiment (maximum efficiency at $pH = 4.8$, the phenol-chloroform solution used being at $pH = 4.1-4.5$). The highly negatively charged RNA molecules always remain in the aqueous phase.

Further extraction with chloroform only aims at reducing the residual concentration of phenol in the aqueous phase. Yet, while the residual phenol may be at a low enough concentration for molecular biology applications, it was still enough to affect our experiments on cells: cells were either killed, or they just greatly slowed down with better removed phenol. This is the reason for steps (i), (iv) and (v). Then the remaining medium was very poor, for most of the proteins and carbohydrate polymers had been depleted, so we diluted it in FM to recover "usual medium richness", hence normal D_{FM} value.

The results presented in figure 3.7 show that this extraction removed the QSF from the HCM_c, while it was still active in the control HCM_c only subjected to the concentration-dilution-concentration cycle. This confirms what was much expected from the estimated molecular weight, namely that the QSF is an amphiphilic large molecule, such as a protein, a carbohydrate or even a lipid aggregate but not an RNA molecule.

The QSF is not degraded by the proteinase K. The proteinase K is an enzyme that digests the proteins by cutting the peptide bonds next to the carboxylic group of hydrophobic and aromatic amino-acids. In other words, it degrades proteins with a broad range and huge efficiency.

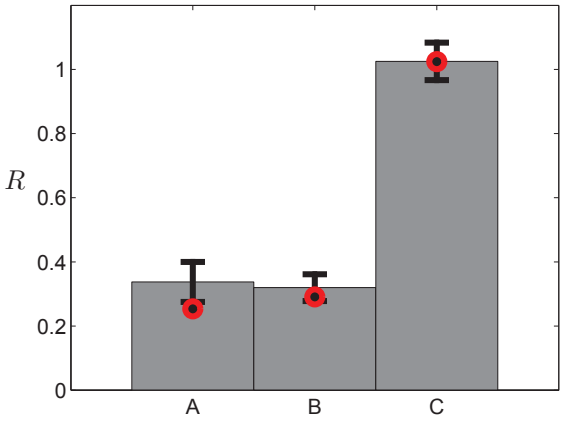


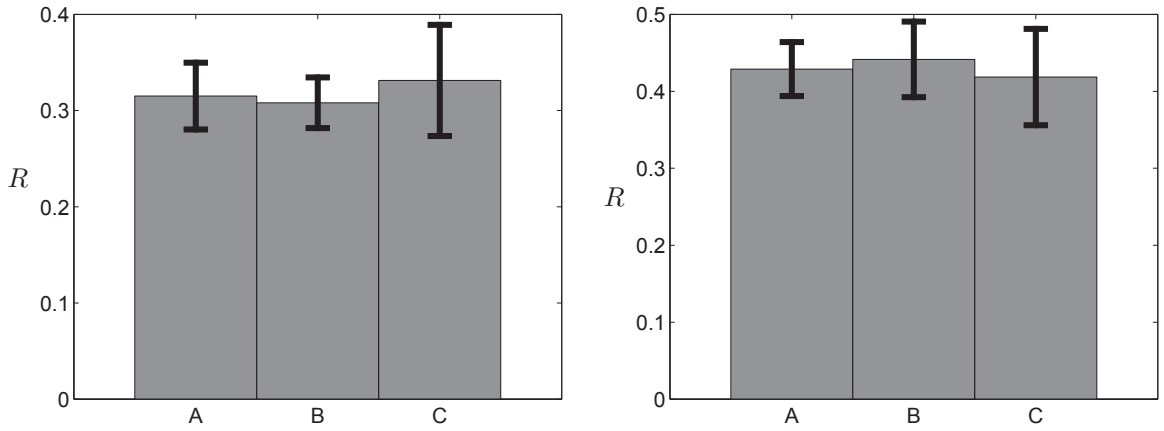
Figure 3.7: QSF stability to phenol-chloroform extraction.
A: control media, $R = \frac{D_{HCM}}{D_{FM}}$ where HCM and FM were not treated.
B: control that the successive concentration/dilution steps do not affect the result. Both FM and HCM were concentrated 20X, diluted 10X, concentrated 20X media, and then diluted 40X in FM.
C: test condition. The media were subjected to the same treatment as in B, with in addition one phenol-chloroform, then two chloroform-only, extractions after the first concentration step.
The red circles correspond to an experiment with 10 times more concentrated treated fraction (last dilution step in B and C, 4X instead of 40X).

Were the QSF to be a protein, treating HCM with proteinase K would very likely suppress the QSF activity.

The usual concentration range for proteinase K is from 10 to 100 $\mu g.mL^{-1}$, with maximum efficiency at 37°C, and it is inactivated above 65°C. As a consequence, in a first experiment, media were incubated with 10 $\mu g.mL^{-1}$ proteinase K (Sigma-Aldrich GmbH, Germany) at 37°C for 2 h, then heated at 65°C for 20 min. In a second one, the concentration was increased to 100 $\mu g.mL^{-1}$ and the proteinase K was eliminated by concentration through a 30 kDa MWCO membrane. Indeed the proteinase K's molecular weight is about 28 kDa.

Neither of the two experiments shows any effect of the treatment by proteinase K on the QSF response. It means that the latter is not likely to be a protein (Fig. 3.8).

The QSF is not charged at experimental pH. FM and HCM were treated with ion exchange resins to eliminate the charged species. As in the previous paragraph, we would expect the QSF response to disappear if the QSF was charged. The cation resin has no effect on motility either



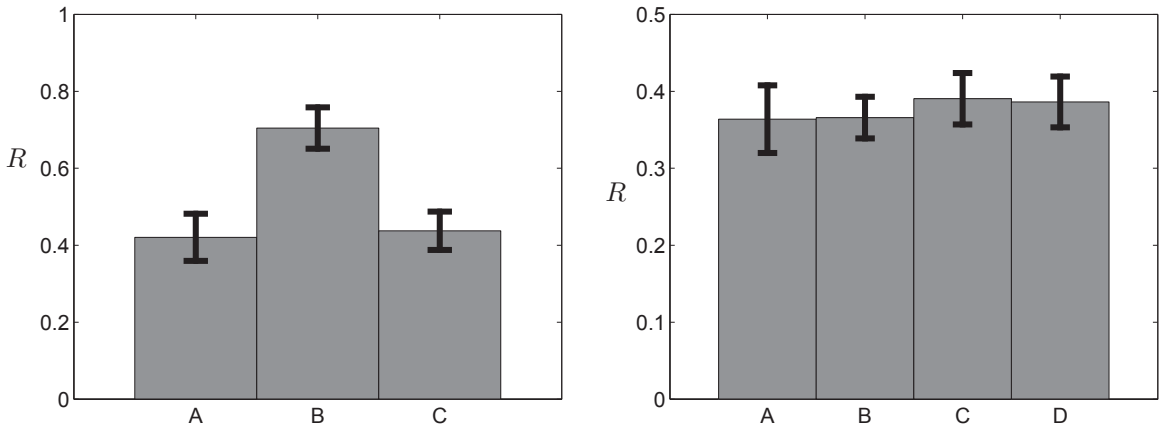
(a) *(b)*

Figure 3.8: QSF stability to moderate heat and protease treatment.
(a) Relative diffusion constant $R = \frac{D_{HCM}}{D_{FM}}$ *in A: control media, B: control for the heat cycle. The media were incubated for* 2 h *at 37°*C*, then for* 20 min *at 65°*C*, C: protease test. Same as B with proteinase K at* 10 µg.mL^{-1}.
(b) QSF response $R = \frac{D_{HCM}}{D_{FM}}$ *in A: control media, B: control for the incubation/filtration steps. The media were incubated for* 2 h *at 37°*C*, then ultra-filtrated through a* 30 k*Da MWCO and mixed in* FM_f*, C: test condition. Same as B with proteinase K at* 100 µg.mL^{-1}.

in FM or in HCM. The anion resin clearly depletes both media of many nutrient, hence yielding increased values for D (Fig. 3.9a). While in simple anion-depleted media R is slightly decreased, in an experiment aiming at complementing the media in nutrients, the full response is recovered (Fig. 3.9b).

It means that the QSF has a low affinity with the resines, hence that it is likely uncharged at pH$= 6.2 - 6.4$.

Taken together, all these characterisation results demonstrate that the QSF is an amphiphilic, uncharged molecule of high molecular weight. If any, its amino-acid part is not predominant as shown by the insensitivity to treatment by the proteinase K and the ion-exchange resins, as peptide sequences are most often charged. The remaining potential candidates include carbohydrate polymers and proteoglycanes among the most probable ones, or even lipid aggregates. In the former case, the QSF would be a linear polymer of molecular weight $M_{QSF} \sim 10$ – 30 kDa, while in the case of an aggregate it could reach the globular weight of $M_{QSF} \sim 50$ – 100 kDa.



(a) *(b)*

Figure 3.9: QSF activity is not sensitive to ionic exchange resine treatment.
(a) Relative diffusion constant R. A: $R = \frac{D_{HCM}}{D_{FM}}$*, B:* $R = \frac{D_{HCM_a}}{D_{FM_a}}$*, C:* $R = \frac{D_{HCM_{cat}}}{D_{FM_{cat}}}$*, where the* $_{cat}$ *(resp.* $_a$*) subscript denotes treatment with a cation (resp. anion) exchange resine.*
(b) Relative diffusion constant $R = \frac{D}{D_{ref}}$*. The 'ref' media are FM (A, B) or* FM_a*:FM (C, D) and the test medium is HCM (A), HCM:FM (B), HCM:*FM_a *(C),* HCM_a*:FM (D). All mixtures are in 1:1 volume proportions.*

3 Cell response and secretion dynamics

During his thesis, Laurent Golé found that the cell response to the QSF was concentration-dependent, and that, for DH1 cells, the response function was correctly fitted by the following expression (equation 3.1):

$$D(c) = D_0 \left(1 + m \exp(-\frac{c}{c_0})\right),$$

with $D_0 = 2\ \mu m^2.min^{-1}$, $m \sim 4$ – 5 and $c_0/\alpha = 3 \times 10^6 cm^2.min^{-1}$, assuming that α, the rate of QSF secretion, is constant. Note that, in this work, the concentrations were measured in cm^{-2} to match with the surface density of cells. Here, we prefer working with volume concentration in mL^{-1}, which is equivalent for a constant medium height and fast homogenisation.

In biological systems, the production and the secretion of chemical factors by the cells are usually under the control of a regulatory network, so that a constant α value is not a very standard situation. To test

these previous findings, we built up our reasoning on the following basic arguments:

$$D = D(c) \tag{3.3}$$
$$\dot{c} = \alpha(\mathbf{p})n(t). \tag{3.4}$$

Equation 3.3 just states that, irrespective of other environmental cues, the cell motility only depends on the QSF concentration c. Equation 3.4 describes the secretion dynamics, where $n(t)$ is the cell concentration at time t. Here we assume a homogeneous surface distribution of the cells and a constant medium height, together with rapid homogenisation of the QSF concentration on this height. The secretion rate depends on a collection of parameters $\mathbf{p}$ through the QSF's particular regulatory network. The simplest situation is when $\alpha = \alpha(c)$ because it involves no other intermediate factor.

Indirect measurement of the QSF secretion. To trace back both the response and secretion dynamics, we followed this procedure:

(*i*) A set of homogeneous cell cultures was prepared at the same initial cell density $n_0 = 3 \times 10^3$ cm^{-2};

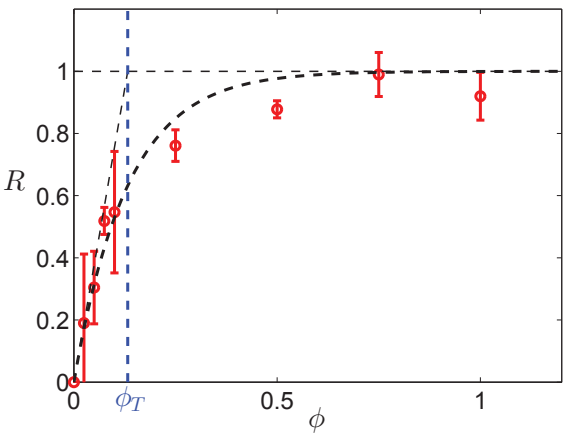
(*ii*) The conditioned media from these dishes were harvested at 10 different times ranging from 4 to 80 hours;

(*iii*) For each of this conditioned media, the cells' diffusion constant was measured at various dilutions in FM, together with a control in FM and a control in a reference HCM.

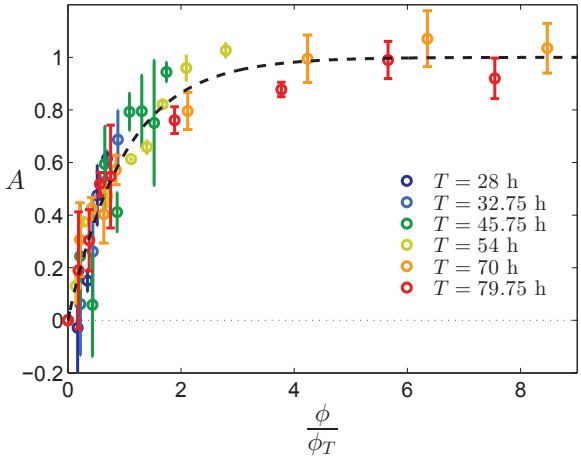
For each of the conditioned media, we obtain a $D(\phi)$ plot, where $\phi = \frac{V}{V+V_{FM}}$ is the volume fraction of the tested medium. Thanks to the two controls, we could convert them into a set of $A(\phi)$ plots using the definition 3.2 to overcome the issue of inter-experiment fluctuations.

The leading idea is that if equation 3.3 is true, then all the curves should collapse onto a single one by the homothety $A = A(\phi.c(T))$ where $c(T)$ is the QSF concentration of the medium harvested at time T (Fig. 3.10).

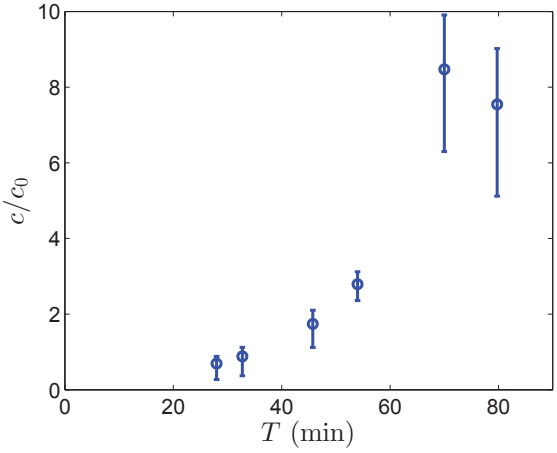
For $T < 22$ h, no QSF activity is measured, even at $\phi = 1$ (no dilution). For $T > 22$ h, we get A functions that increase with ϕ, a later T yielding a steeper increase, with an apparent upper bound at $A = 1$. To collapse the data, we used L. Golé's empiric response function 3.1, which worked



(a)



(b)



(c)

Figure 3.10: Indirect measurement of the QSF concentration. (a) Activity as a function of the dilution for the media harvested at $T = 79.75$ h. We deduce the value of ϕ_T from a fit with $A(\phi) = 1 - \exp(-\phi/\phi_T)$ (b) Collapse of the activities as a function of ϕ/ϕ_T for media collected at various culture times T. (c) Normalised QSF concentration in the conditioned media ($c(T)/c_0 = 1/\phi_T$), as a function of the culture time T.

very satisfactorily. Notice that in terms of activity, this function has the advantage of having one single parameter, the QSF response threshold c_0. Indeed:

$$A(c) = \frac{D_{FM} - D(c)}{D_{FM} - D_{HCM}}$$

$$= \frac{D_0.(1+m) - D_0.(1 + m\exp(-c/c_0))}{D_0.(1+m) - D_0}$$

$$= 1 - \exp\left(-\frac{c}{c_0}\right). \tag{3.5}$$

Now, in terms of dilution ϕ, which is our work variable, rather than in terms of concentration, which is yet to be determined, we get:

$$\begin{aligned} A(c) &= 1 - \exp\left(-\frac{c}{c_0}\right) \\ &= 1 - \exp\left(-\frac{\phi.c(T)}{c_0}\right) \\ &= 1 - \exp\left(-\frac{\phi}{\phi_T}\right). \end{aligned} \tag{3.6}$$

The fitting parameter being $\phi_T = \frac{c_0}{c(T)}$, we then get the QSF concentration at various times T up to a factor: $c(T) \sim \frac{1}{\phi_T}$.

The results of these measurements are displayed in figure figure 3.10b. The QSF concentration seems to saturate, for $c(70\text{ h}) \simeq c(80\text{ h})$. Clearly, this cannot be accounted for by a constant secretion rate α. More precisely, it appears that α decreases in time. To estimate its evolution, we assumed it constant between two consecutive data-points. In that approximation, the secretion dynamics can be solved for each time window, knowing that the cell proliferation that was measured in parallel has an exponential behaviour (Fig. 3.11a), as expected:

$$n(t) = n_0 2^{\frac{t}{t_2}}, \tag{3.7}$$

with a doubling time $t_2 = 8.5 \pm 0.3$ h.

Then, equation 3.4 yields, where α_i denotes the average α value between T_{i-1} and T_i:

$$\int_{c(T_{i-1})}^{c(T_i)} \mathrm{d}c = \alpha_i \int_{T_{i-1}}^{T_i} n(t)\mathrm{d}t$$

$$\Rightarrow c(T_i) - c(T_{i-1}) = \alpha_i \frac{n_0 t_2}{\ln 2}\left(2^{\frac{T_{i-1}}{t_2}} - 2^{\frac{T_i}{t_2}}\right)$$

$$\Leftrightarrow \alpha_i = \frac{\ln 2\,[c(T_i) - c(T_{i-1})]}{t_2\,[n(T_i) - n(T_{i-1})]}. \tag{3.8}$$

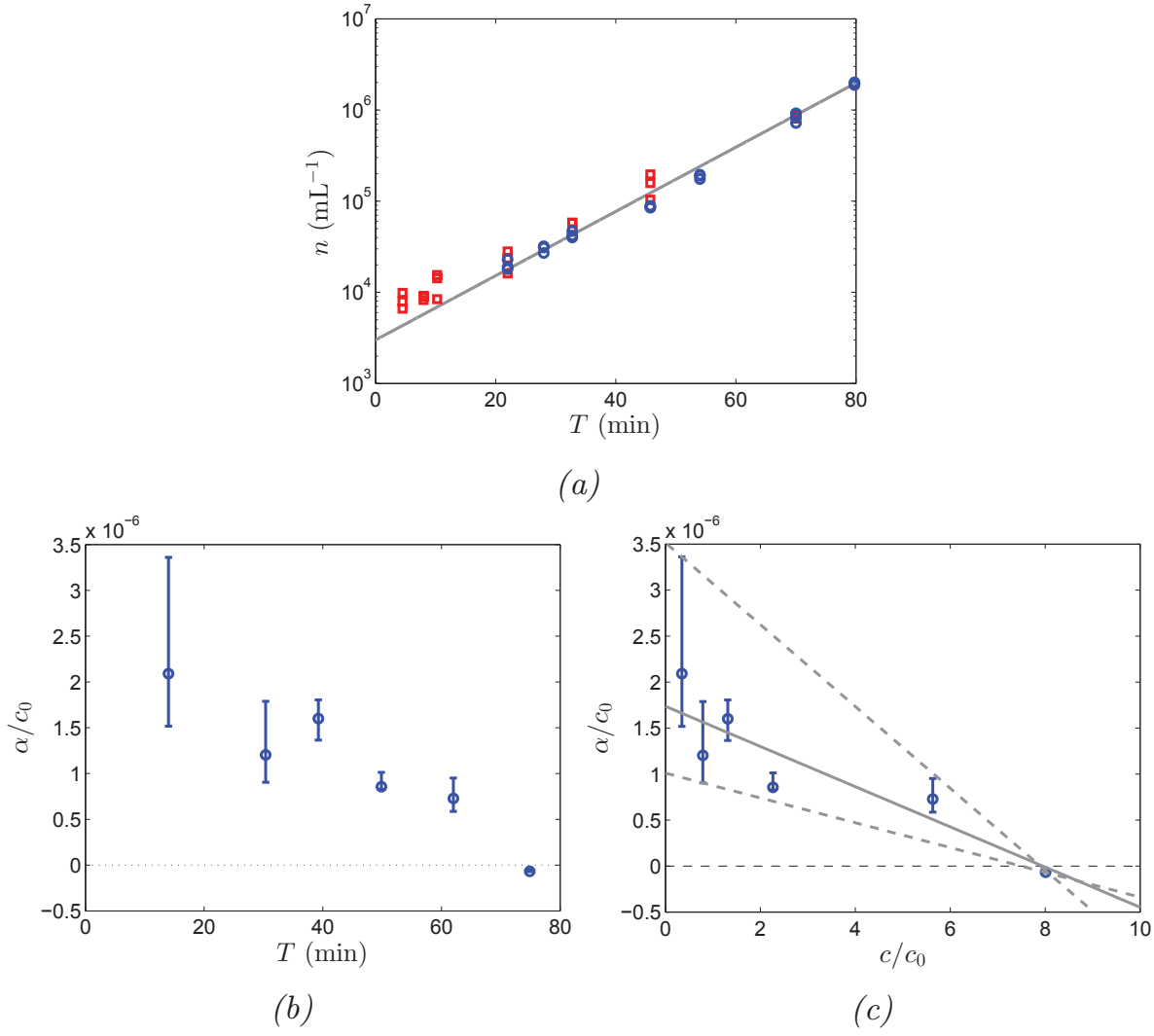


Figure 3.11: (a) Cell proliferation, as measured from pictures of the culture dishes (red squares) or manual count with a Malassez slide (blue circles). The grey line is an exponential fit with $t_2 = 8.5$ h. (b) Normalised secretion rate computed from $c(T)$ and $n(T)$ as a function of T. (c) Normalised secretion rate as a function of c. The solid line is a linear fit, while the dashed lines illustrate the parameter variability (limit cases that include all the error bars). The fit yields $\frac{\alpha_0}{c_0} = 1.7 \times 10^{-6}$ mL.h^{-1} and $\frac{c_m}{c_0} = 8.0$ while the extreme intervals are $\frac{\alpha_0}{c_0} \in [1.0 - 3.5] \times 10^{-6}$ mL.h^{-1} and $\frac{c_m}{c_0} \in [7.5 - 8.0]$.

The computed values of α_i are plotted against those of c_i in figure 3.11c. It shows that the data are compatible with a decreasing $\alpha(c)$ relationship, with a fixed point at $c = c_m$. The simplest possible shape for this relationship, which seems quite reasonable, is an affine one:

$$\alpha(c) = \alpha_0 \left(1 - \frac{c}{c_m}\right). \tag{3.9}$$

This yields a linear secretion equation that can again be solved analytically:

$$\frac{\mathrm{d}c}{\alpha_0\left(1-\frac{c}{c_m}\right)} = n(t)\mathrm{d}t. \tag{3.10}$$

The right-hand side is the same as before, hence by integration:

$$\int_0^t n(t)\mathrm{d}t = \frac{n_0 t_2}{\ln 2}\left(2^{\frac{t}{t_2}}-1\right). \tag{3.11}$$

To integrate the left-hand side, we use the change of variable $u = c_m - c$ to get:

$$\begin{aligned}\frac{c_m}{\alpha_0}\int_0^{c(t)}\frac{\mathrm{d}c'}{c_m-c'} &= \frac{c_m}{\alpha_0}\int_{c_m-c(t)}^{c_m}\frac{du}{u}\\ &= \frac{c_m}{\alpha_0}\left[\ln(c_m)-\ln\left(c_m-c(t)\right)\right].\end{aligned} \tag{3.12}$$

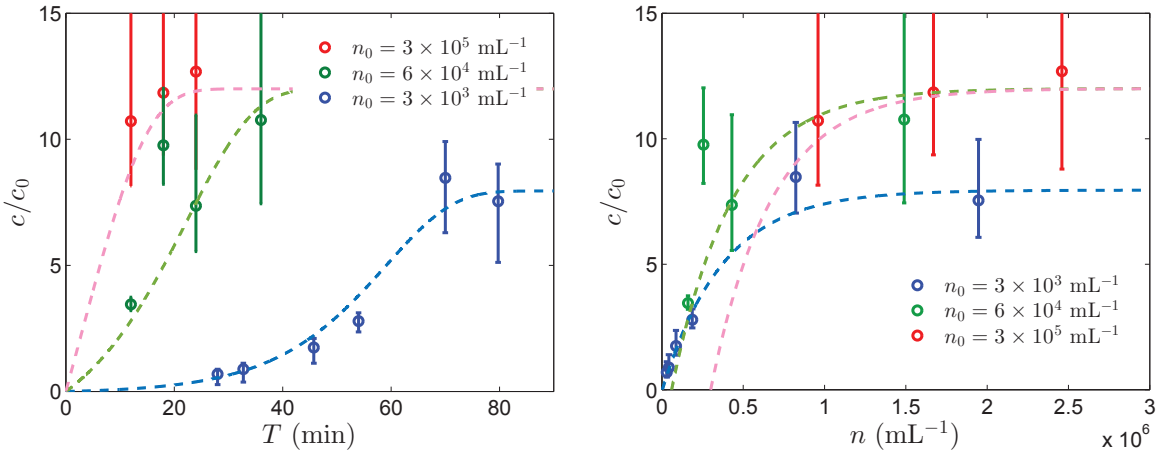
Equating both integrated sides of 3.10, we obtain:

$$\begin{aligned}c(t) &= c_m - \exp\left[\ln(c_m)-\frac{\alpha_0 n_0 t_2}{c_m\ln 2}\left(2^{\frac{t}{t_2}}-1\right)\right]\\ &= c_m\left(1-\exp\left[-\frac{\alpha_0 n_0 t_2}{c_m\ln 2}\left(2^{\frac{t}{t_2}}-1\right)\right]\right).\end{aligned} \tag{3.13}$$

Using this expression with $\frac{\alpha_0}{c_0} = 1.7\times10^{-6}$ mL.h^{-1} and $\frac{c}{c_0} = 8.0$ fits our measurements very correctly. To confirm this finding of a negative feedback on the QSF secretion, we did the same experiment as before, starting from two different cell densities: $n_1 = 6\times10^4$ mL^{-1} and $n_2 = 3\times10^5$ mL^{-1}. In the case of constant α, the QSF concentration at a given t should be respectively twenty and one hundred times higher than in the previous experiments with $n_0 = 3\times10^3$ mL^{-1}, whereas our

linearly decreasing $\alpha(c)$ would yield values of c next to the saturation for the harvest times $T = 12$, 18, 24, and 36 h.

This second experiment was done with less accuracy than the previous one – one realisation for each condition instead of three – and more importantly, the reference HCM was not the same as before. Yet, the $A(\phi)$ curves allowed to retrieve again $c(t)$ up to a factor c_0. As expected they seem close to a saturation value c_m, but c_m seems *a priori* different from that of the previous experiment (Fig. 3.12). However, using a different reference HCM could have been misleading for the estimation of c_0, hence of c_m. Rescaling α_0 so that $\frac{\alpha_0}{c_m} = 2.2 \times 10^{-7}$ mL.h^{-1} is the same for both experiments makes both sets of data compatible through the use of 3.13.



(a) *(b)*

Figure 3.12: Comparison of the analytical results to the experiments. (a) Normalised QSF concentration as a function of time for three different initial cell densities: $n_0 = 3 \times 10^3$ mL^{-1} (blue), $n_0 = 6 \times 10^4$ mL^{-1} (green) and $n_0 = 3 \times 10^5$ mL^{-1} (red). The dashed lines are plots of the expression 3.13 with the parameter values from figure 3.11 (b) Normalised QSF concentration as a function of n for three n_0. The dashed lines are plots of the expression 3.14

A robust self-counting system. When it comes to measuring the current cell density, the drawback of a constantly secreted counting factor is that its concentration is the integral of the cell density. In other words, it results from the accumulation of factor both in the present and in the past, hence being not only dependent on the current cell density, but also very sensitive to the precise history of the cell population. As such, it is not an efficient system to estimate the cell density. Having a feedback on the secretion dynamics could possibly solve this issue, and we wanted to

test how much the linearly decreasing $\alpha(c)$ that we found could improve this.

Using the expressions 3.7 and 3.13, we computed both the cell density $n(t)$ and the QSF concentration $c(t)$, starting from various initial cell densities. The results displayed in figure 3.13 show that $c(n)$ does not depend much on the initial condition for $n \gg n_0$. In particular, in this approximation, $c(n)$ rewrites:

$$c = c_m \left(1 - \exp\left[-\frac{\alpha_0 t_2}{c_m \ln 2}\left(n(t) - n_0\right)\right]\right) \tag{3.14}$$

$$\approx c_m \left(1 - \exp\left[-\frac{\alpha_0 t_2}{c_m \ln 2} n\right]\right). \tag{3.15}$$

This appears to be an efficient cell counting mechanism. Indeed, although its range is not very broad – c equals either 0 or c_m for almost all cell densities – it is not sensitive to past history, and it may allow to selectively detect cell densities around $n^* = \frac{c_m \ln 2}{\alpha_0 t_2}$, as illustrated in Figure 3.13. Indeed, c reaches 63% of its saturation value at $n = n^*$, and 95% at $n \simeq 3n^*$. Interestingly, the obtained value of $n^* = 3.7 \times 10^5$ mL^{-1} is close to the maximum value that is empirically recommended for *Dictyostelium* cultures.

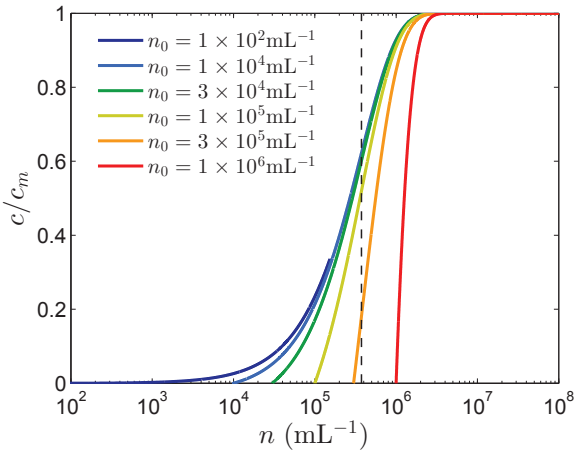


Figure 3.13: Theoretical QSF concentration using expression 3.14 as a function of n for various initial cell densities from 1×10^2 mL^{-1} *to* 1×10^6 mL^{-1}. *The vertical dashed line is located at* $n = n^*$.

Thus, using measurements of the QSF activity in dilutions of various conditioned media with respect to a reference FM and a reference HCM, we first showed that the empirical expression (3.1), originally proposed

to account for the QSF response of DH1 cells, describes well the cellular response to the QSF concentration, even in AX2 cells. Then, taking advantage of this robust response function, we could retrieve the value of the QSF concentration in all those conditioned media with a single parameter $\frac{\alpha}{c_m}$, up to a factor. We found that the secretion rate is not constant, but that the QSF concentration rather exhibits a saturation value c_m. Our data are compatible with a linearly decreasing secretion rate, which allows the analytical calculation of the QSF concentration in time. This may constitute a robust quorum-sensing system with high selectivity in cell density. Generally speaking, such a mechanism could be used to perform any function that should be triggered by a given cell density, the latter being set by the respective values of the cells' doubling time, the secretion rate and the saturation concentration.

4 A journey into response pathways

In order for the QSF concentration to be translated in terms of motility, cells have to process the signal through a mechanism called the response pathway, or signalisation cascade. It consists in a series of molecular steps of various natures: chemical intermediaries such as proteins, RNAs, nucleotides, lipids, carbohydrates or other small molecules can be activated, deactivated, produced or destroyed by catalysis, chemical or structural modifications.

In order for the signal to cross the membrane, the QSF can be either bound by a membrane receptor or internalised by endocytosis for instance. Then the signal is processed through the cascade, which can be linear but can also involve many ramifications and feedback loops, to finally activate a response that can be of three major natures:

(*i*) genetic regulation: the signal is imported into the nucleus where it acts to regulate the transcription of a gene into mRNA;

(*ii*) translational regulation: the signal acts to regulate the translation of a mRNA into a protein;

(*iii*) direct regulation: the signal directly acts on the phenotypic activity – *e.g.* on the acto-myosin activity in the case of motility.

This cascade is very important for two main reasons: firstly, it defines molecules that are more or less important for the response to happen. For instance, if a protein constitutes a bottleneck in the response pathway,

a mutant that does not express it will not exhibit any response to the QSF. Secondly, the response function and secretion feedback that we studied mathematically in the previous section depend completely on its spatio-temporal dynamics.

As a consequence, we conducted a series of experiment aiming at determining whether some proteins were involved in the QSF signalisation pathway, so as to sketch its contours.

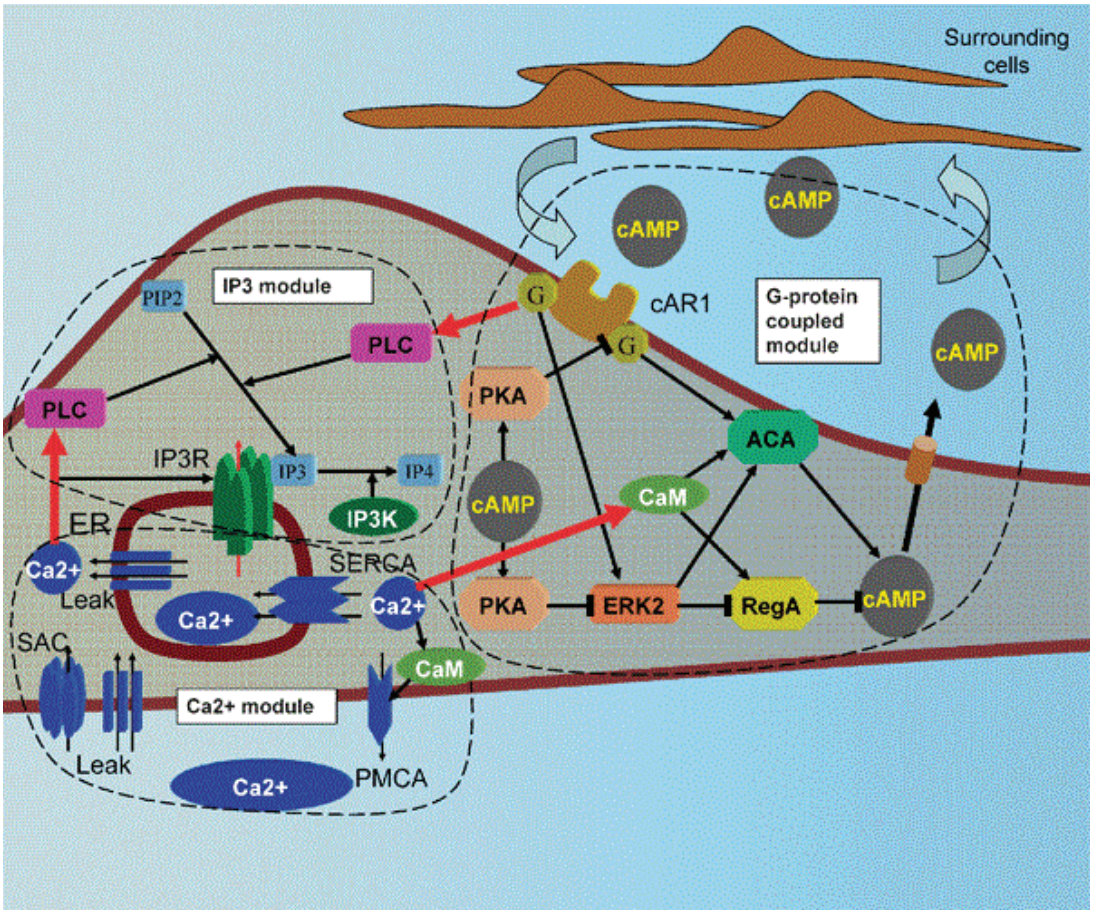


Figure 3.14: Dictyostelium's signalling pathways (from [90]).
The three main paths are known as the "phosphate pathway" (top left), which involves in particular the PhosphoLipase C (PlC) under the control of a G-protein complex, the "calcium pathway" and the "cAMP pathway". The latter is related to external signals through receptor-binding G-proteins that activate the Adenylate Cyclase A (ACA). The latter induces the production of cAMP, which regulates its production by triggering a feedback loop initiated by the Protein Kinase A (PKA).

4.1 Involvement of receptor-binding G-protein subunits

First we wanted to ascertain how the cells sense the QSF. In most cases it is realised by transmembrane receptor proteins or protein complexes, which change conformation upon ligand binding, hence initiating a signalisation cascade inside the cytoplasm.

Dictyostelium discoideum has many receptors, some of which are not well characterised or does not have a null mutant. Thus, trying to find directly the receptor would have been quite hazardous. Yet, *Dictyostelium* posses an interesting receptor-bound G-protein complex, which constitutes a kind of bottleneck for many receptor-mediated signalling pathways (Fig. 3.14). This complex is made up of three subunits, called Gα, Gβ and Gγ– also called gpa, gpb and gpg in the*Dictyostelium* systematic notation. While Gα has 12 identified variants, Gβ has only two and Gγ only one. We could test only a subset of these proteins, as all null mutants were not available from the Dicty stock center (`http://www.dictybase.org/`), a NIH ruled institute that maintains all referenced *Dictyostelium* strains.

We measured the diffusion constant of null mutants of these G-protein subunits to determine whether they were involved or not. A first experiment with $G\beta 1^-$ cells showed that in this strain, the relative diffusion constant R is quite close to 1 (Fig. 3.15) . Doing again this experiment (Fig. 3.16a), we found again a partial suppression of the response, with surprisingly spread values of R this time. This advocates the idea that the response is mediated by a G-protein-binding membrane receptor, but that Gβ1 can be bypassed in some way.

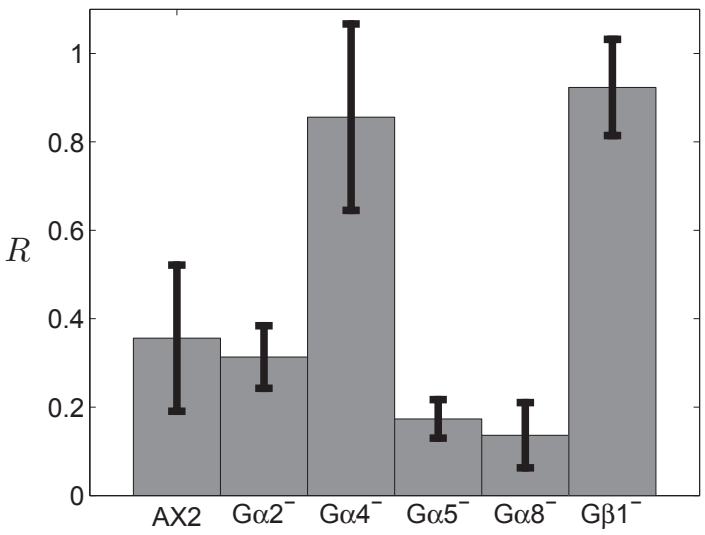
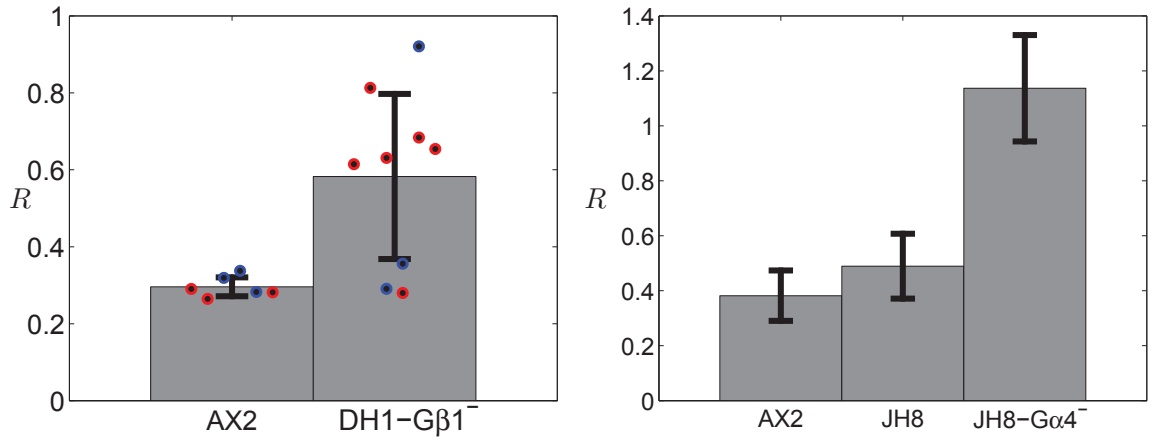


Figure 3.15: Involvement of a receptor-binding G-protein complex.
QSF response for various Ga and Gβ knock-out cells. $Ga4^-$ and $G\beta 1^-$ cells exhibit a very limited response to the QSF – ie relative diffusion constant R close to 1.

Then, we tested the motility of a series of mutants of various Gα subunits: $G\alpha2^-$, $G\alpha5^-$ and $G\alpha8^-$ mutants were found to respond normally to the QSF, while $G\alpha4^-$ cells did not (Fig. 3.15). The last strain stemmed from the JH8 strain, which we also tested in a second experiment (Fig. 3.16b). The result is that the parental strain JH8 has a low motility compared to AX2, but that the response to the QSF – secreted by AX2 cells – is still observed with a similar R. In contrast, JH8-$G\alpha4^-$ cells have a similar motility as JH8 in FM, but they keep the same motility in HCM ($R \simeq 1$), hence not responding to the QSF.

From these results, we conclude that the response to the QSF is mediated by a membrane receptor and involves a receptor-binding G-protein complex, where the Gα4 subunit is essential for the response and the Gβ1 subunit is involved but dispensable.



(a) *(b)*

Figure 3.16: Repeat of the experiment with the identified G-protein subunits.
(a) Relative diffusion constant R for control AX2 cells and DH1-Gβ1⁻ cells. Mean and standard deviation from two colour-coded series of 3 or 6 independent experiments. The disks show the individual results from all independent experiments, which exhibit a wide spreading in the case of Gβ1- cells with both usual and notably higher than usual values of R. (b)Relative diffusion constant for control AX2, JH8 and JH8-Gα4⁻ cells in a second series of experiments. The Gα4⁻ cells do not respond to the QSF while the parental JH8 cells do.

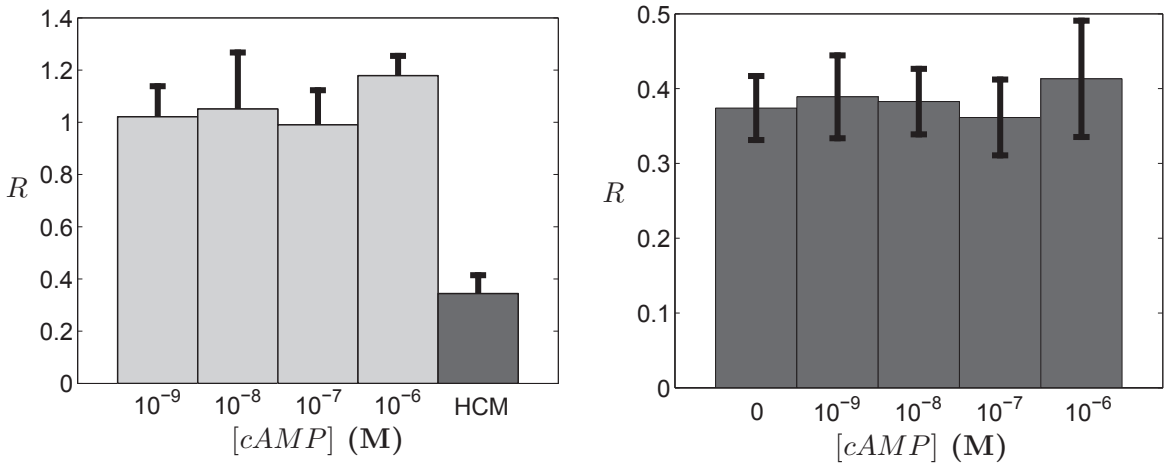
4.2 The role of cyclic AMP and known signalling pathways.

A well-known secreted molecule that acts to regulate the motility of *Dictyostelium discoideum* cells is the cyclic adenosine monophosphate

(hereafter simply called cAMP). This cyclic nucleotide is secreted in a pulsatile manner upon starvation, which gives rise to concentration waves through signal relay and cAMP consumption, and is eventually responsible for the cell streaming and aggregation. It acts on the motility of starved cells both by polarising the cells in the direction of the wave's source – or alternatively upwards in a simple cAMP gradient – and by increasing their migration activity. Moreover, it does not only play a role as a secreted molecule, but also as a cytoplasmic signal, which makes its homoeostasis quite complicated, given that external and internal cAMP have different functions.

External cAMP. First, we tested whether cAMP could act as an external messenger in the QSF regulation, either that it be the QSF itself – although it is very unlikely given the biochemical characterisation results, cAMP is a too important secreted signal in *Dictyostelium* to afford not to test it – or an interfering co-messenger.

To that end, we measured the QSF response of AX2 cells in FM supplemented with various concentrations of cAMP (see Fig. 3.17a) or HCM treated in the same way (Fig. 3.17b). It is quite clear that over a wide range of cAMP concentrations, which correspond either to the physiological one or those used *in vitro*, *e.g.* for chemotaxis assays, cAMP has no influence on the cells' diffusion constant, and especially it does not modify the relative diffusion constant R.



(a) *(b)*

Figure 3.17: Adding cAMP in the medium does not affect the cells' motility.
(a) Relative diffusion constant R for cAMP supplemented FM at various concentrations [cAMP] and control in pure HCM. (b) Relative diffusion constant R for HCM supplemented with various concentrations [cAMP].

Internal cAMP. Then, to check whether an internal production of cAMP could be involved in the response to the QSF, we used the cAMP analogue 8-Br-cAMP. When it is present in the medium, this molecule is internalised, contrary to cAMP, and it activates the same targets as cAMP. Therefore, adding some 8-Br-cAMP in the medium is similar to forcing the production of cytoplasmic cAMP. Would the addition of 8-Br-cAMP in HCM suppress the QSF response, it would mean that internal cAMP does so, and so lead to the hypothesis that in the presence of QSF, the internal cAMP production is inhibited. On the contrary, a significant decrease of D upon supplementation of FM with 8-Br-cAMP could mean that the QSF promotes the production of cAMP.

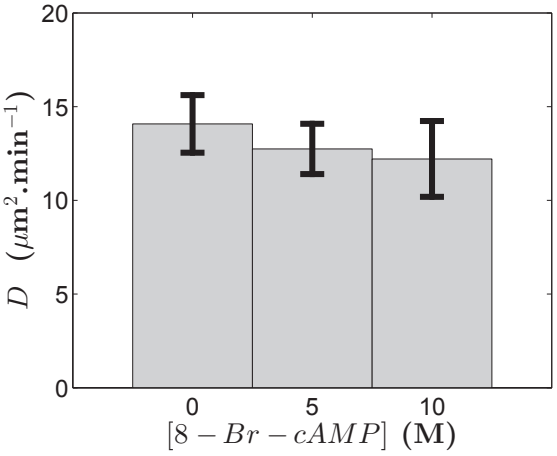
Looking at D for AX2 cells in FM with or without 8-Br-cAMP, and R for the same cells in HCM submitted to the same treatment, this internal cAMP analogue seems to have very little, if not no effect (Fig. 3.18a,b). Yet it is known that 8-Br-cAMP takes about 30 – 45 min to penetrate the cells' cytoplasm and activate the pathway. Considering shorter trajectories that starts at $t = 45$ min does not change the results a lot – though the obtained values of D are slightly lower, since we get closer to the persistence time. Looking at the evolution of D over hours for one longer experiments confirms that 8-Br-cAMP has no effect on HCM. It seems to have some effect on FM after a time, but it is hard to conclude, firstly because the effect is not very strong, but more importantly because it is not monotonous with the 8-Br-cAMP concentration (see Fig. 3.18c).

From these two series of experiments, we conclude that neither external nor internal cAMP seem to play a fundamental role in the QSF regulation of the cellular motion. In what follows, we explore further the cAMP signalling pathway, together with other classical ones.

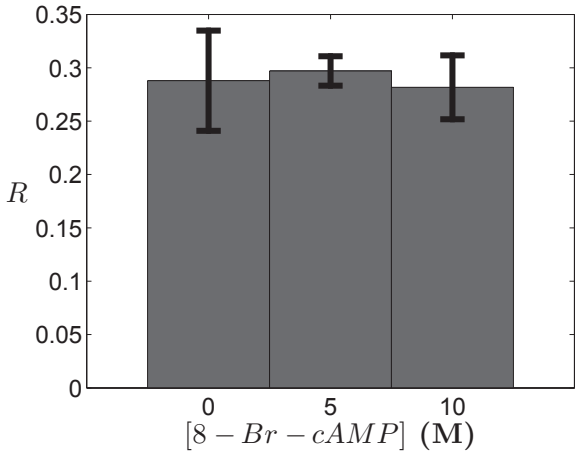
4.3 A trip to hot spots of *Dictyostelium*'s signalling.

Indeed, as we stated above, cAMP behaves not only a secreted messenger but it also plays a role in intracellular signalisation, through the so-called cAMP pathway, which involves the activation of ACA by a $G\alpha\beta\gamma$ complex and a subsequent production of cAMP. This in turn activates PKA through the release of the PKAC subunit, which introduces a negative feedback by inhibiting both the $G\alpha\beta\gamma$ complex and the cAMP production by ACA, while favouring the cAMP hydrolyse by RegA under the negative control of ERK2 (see Figure 3.14).

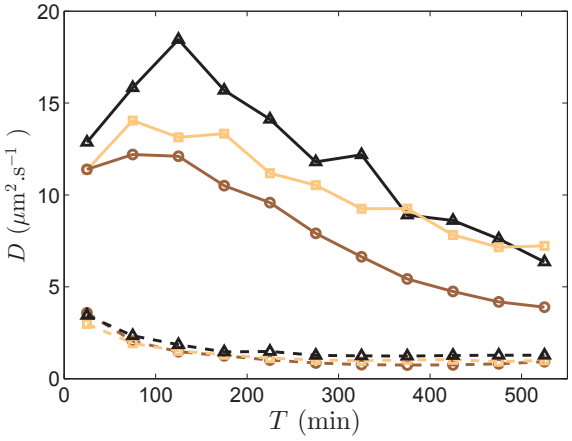
Two other major pathways that are linked to the regulation of the motility in *Dictyostelium* are the calcium pathway and the phosphate pathway



(a)



(b)



(c)

Figure 3.18: Effect of internal cAMP stimulation.
(a) Diffusion constant of AX2 cells in FM supplemented with 8-Br-cAMP at different concentrations. (b) Relative diffusion constant $R = \frac{D_{HCM}}{D_{FM}}$ *of AX2 cells in HCM and FM supplemented with 8-Br-cAMP at different concentration (the same concentration both in FM and HCM for each point). (c) Time evolution of D in FM (solid lines) and HCM (dashed lines) with* 0, 5 *and* 10 m*M of 8-Br-cAMP (resp. black up triangles, brown circles and ochre squares).*

(Fig. 3.14). Superoxide signalling is also known to be necessary for the early multicellular development phase of starved *Dictyotselium* cells [91]. The last was evoked in the introduction section: it is the AprA regulatory system, which involves the co-secreted protein CfaD, the receptor binding Gα8 protein and the cytoplasmic kinase QkgA for both its activities, namely the regulation of the proliferation and chemorepulsion [11, 22].

In the present section, we explore these various pathways by studying the QSF response of mutants that lack important intermediaries of these pathways.

AprA. We first checked whether the QSF was AprA itself. To that end, we measured the response of $aprA^-$ mutants to the QSF, and

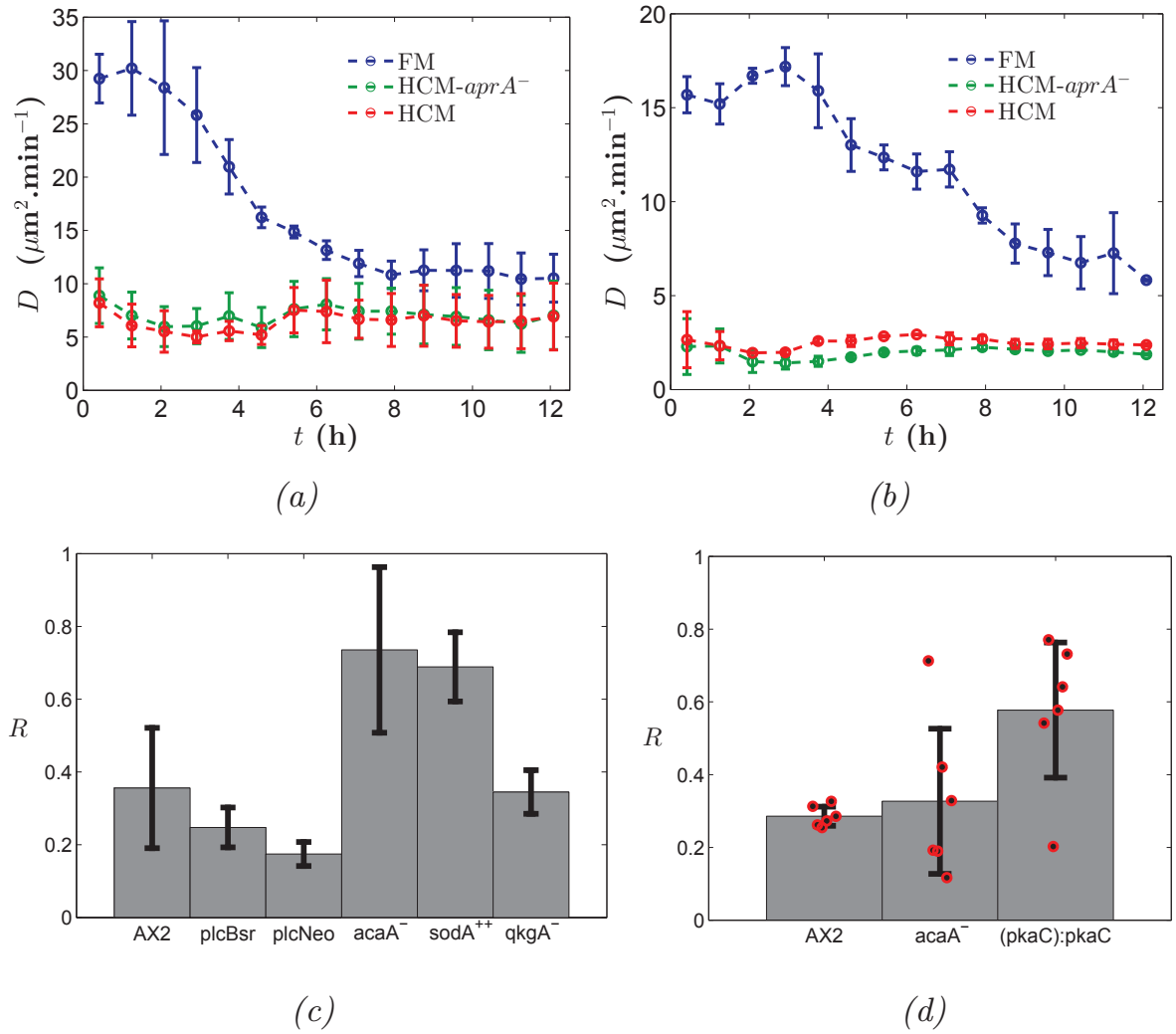


Figure 3.19: Response of various mutants to the QSF.
(a) Time evolution of D for AX2 cells in FM (blue), HCM (red) and HCM made by aprA⁻ cells (green). The conditioning of the FM with time is clear, as well as the initial conditioning of HCM even in the absence of AprA. (b) Time evolution of D for AX2-aprA⁻ cells. The media are the same as in (a). It is clear that aprA⁻ cells respond to the QSF in both HCM, and that they condition the FM with time. (c) Relative diffusion constant R of various mutants. The response looks weaker but not suppressed in acaA⁻ and (act15):sodA cells (see main text for more details). (d) Relative diffusion constant R of acaA⁻ and (pkaC):pkaC) cells in a later experiments. Both strains respond to the QSF, even though the response is decreased, with widespread values, in the case of the faster (pkaC):pkaC cells. Red circles denote the values obtained from individual independent experiments.

showed that *aprA⁻* are perfectly able to self condition their medium (Fig. 3.19b). We also measured the response of wild-type cells to HCM

made from $aprA^-$ cultures (Fig. 3.19a), HCM from AX2 treated with an anti-AprA antibody, and FM supplemented with recombinant AprA (rAprA), kindly contributed by Richard Gomer (Department of Biology, Texas A&M University). All these experiments suggest that AprA is not the searched QSF, even though a slight decrease of the motility is induced in a certain range of rAprA concentrations.

QkgA. The ROCO kinase QkgA is necessary for both activities of secreted AprA. Consistently with the results on AprA, it is not involved in the QSF response, as shown in Figure 3.19c.

AcaA. The adenylate cyclase of aggregation stage, or adenylyl cyclase AcaA is responsible for the production of cAMP, under the control of G-protein complexes. In a first experiment (Fig. 3.19c), $acaA^-$ cells seemed to exhibit R values closer to 1, while in a more recent one (Fig. 3.19d) it was not the case. However, these results are very hard to interpret correctly as in both cases a substantial fraction of cells did not adhere to the substrate.

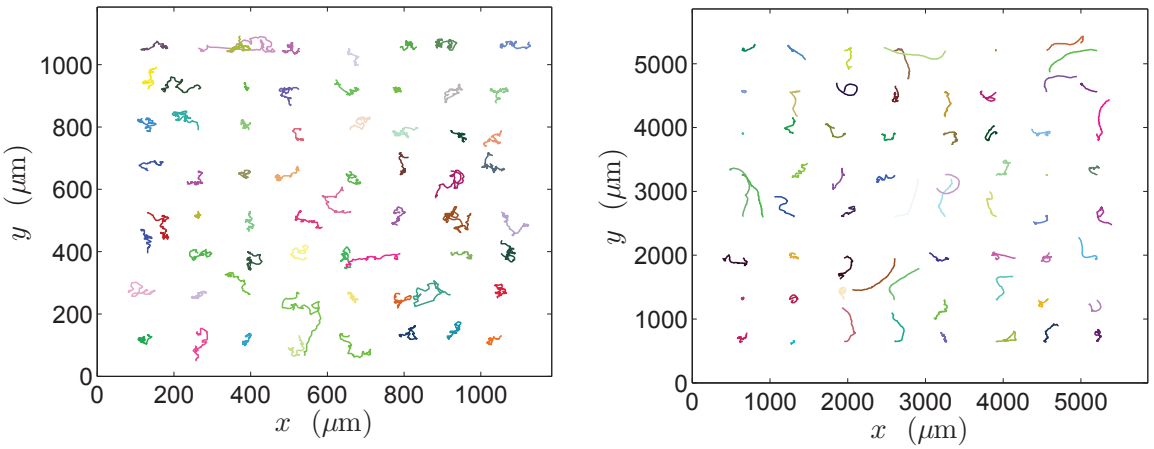
PkaC. The catalytic subunit of protein kinase A PkaC introduces a negative feedback in the cAMP pathway: indeed, it is activated by the cAMP, but it provokes a decrease of the cAMP concentration both by down-regulating its production and by promoting its hydrolysis by RegA. AX2 cells that over-express PkaC, noted $(pkaC) : pkaC$ cells, exhibit an increased D both in FM and in HCM. This is quite unexpected, as a higher level of PkaC lowers intracellular cAMP, which should decrease the motility.

However, there is still a response to the QSF, which is more relevant to the present work, even though it is reduced, confirming that the cAMP pathway might well slightly interfere with the QSF response (Fig. 3.19d).

Plc. The phosphoinositide-specific phospholipase C, noted Plc or alternatively PipA, intervenes in the phosphate pathway and it is especially necessary for normal chemotaxis. We checked that two different plc^- knock-outs responded normally to the QSF, showing that this pathway is likely not involved in the QSF regulation of the motility (Fig. 3.19c)

SodA. Lastly, we measured the response of *(act15):sodA* cells that overexpress the superoxide dismutase, which catalyses the dismutation

of superoxide ions, hence reducing the concentration of free radicals. Superoxide signalling has been shown to be important in chemotaxis, with SodA overexpression leading to the suppression of aggregation for instance [91].



(a) (b)

Figure 3.20: (act15):sodA cells exhibit smooth trajectories. Trajectories of (a) AX2 and (b) (act15):sodA cells in FM. The origins of the trajectories have been put on a regular array for the sake of clarity. Warning: the scales in (a) and (b) are different.

In our experiments, *(act15):sodA* cells exhibit a reduced response to the QSF (Fig. 3.19c). However, they also have a particularly high diffusion constant, and their trajectories look very dissimilar from all the other one: they are much smoother both in FM and HCM (Fig. 3.20). From that, we hypothesise that SodA overexpression disturbs the motility by an other means than the QSF. Although it would probably be very interesting to look further into these particular trajectories, or alternatively to take advantage of them as a parameter button in other experiments, this pathway does not appear to be relevant to the QSF regulatory system.

4.4 Summary

We started addressing the issue of the signalling pathway involved in the response to the QSF. We found that the response is mediated by receptor-binding G-protein subunits. In particular, the Gα4 subunit, previously shown to be involved in folate chemotaxis and morphogenesis [92, 93], is necessary for the response.

Then we stepped into the main signalling pathways of *Dictyostelium discoideum* (Fig. 3.14). The AprA/CfaD system does not seem to be

involved, as well as the phosphate path. Superoxide signalling interferes with the response to the QSF but the motility of SodA over-expressing cells is too different from the wild-type to draw a definitive conclusion from this experiment. Last, the cAMP pathway, which is so fundamental in *Dictyosetlium*'s motility, is apparently not directly involved. However, PkaC over-expressing cells show a reduced response. Whether it is due to actual involvement of PkaC in the QSF response or not, again remains unclear, but this way should not be abandoned right away as the MAP kinase Erk2, closely related to PkaC, is known to interact with the Gα4 subunit [93].

5 How does the QSF affect single trajectories?

Even though, in terms of signalling, the cellular response to the QSF remains elusive to that point, we can make use of the rich trajectory data that we have to study the response at the scale of single cells' motion.

In his previous work on DH1 cells [18], Laurent Golé claimed that the regulation was made at the level of the persistence time, with a "constant speed" emerging from the linear relation between D and P, the persistence time extracted from a fit of $MSD(\delta t)$ using the Fürth formula. However, we have seen that this formula may not apply well to the considered trajectories, and especially to measure the persistence time in the case of DH1 cells that exhibit a short-time negative correlation. Moreover, this assertion was not supported by any measurement of the speed, even though we have seen that it might not be well-defined.

In this section, we start tackling this issue based on the numerous data obtained from the FM and HCM controls of 156 experiments of this chapter, with a total of 25,858 trajectories in FM and 28,056 trajectories in HCM.

Variability in the diffusion constant. Let us first draw a general picture of the dataset that we have. It consists of 156 sets of trajectories of cells in FM, and the corresponding ones from cells in HCM recorded *in the same experiment*. This last point is very important, to the extent that it allows us to distinguish the effects of inter-experimental variability

due to environmental or cell-culture-related non-controlled parameters, from those actually stemming from the presence of QSF.

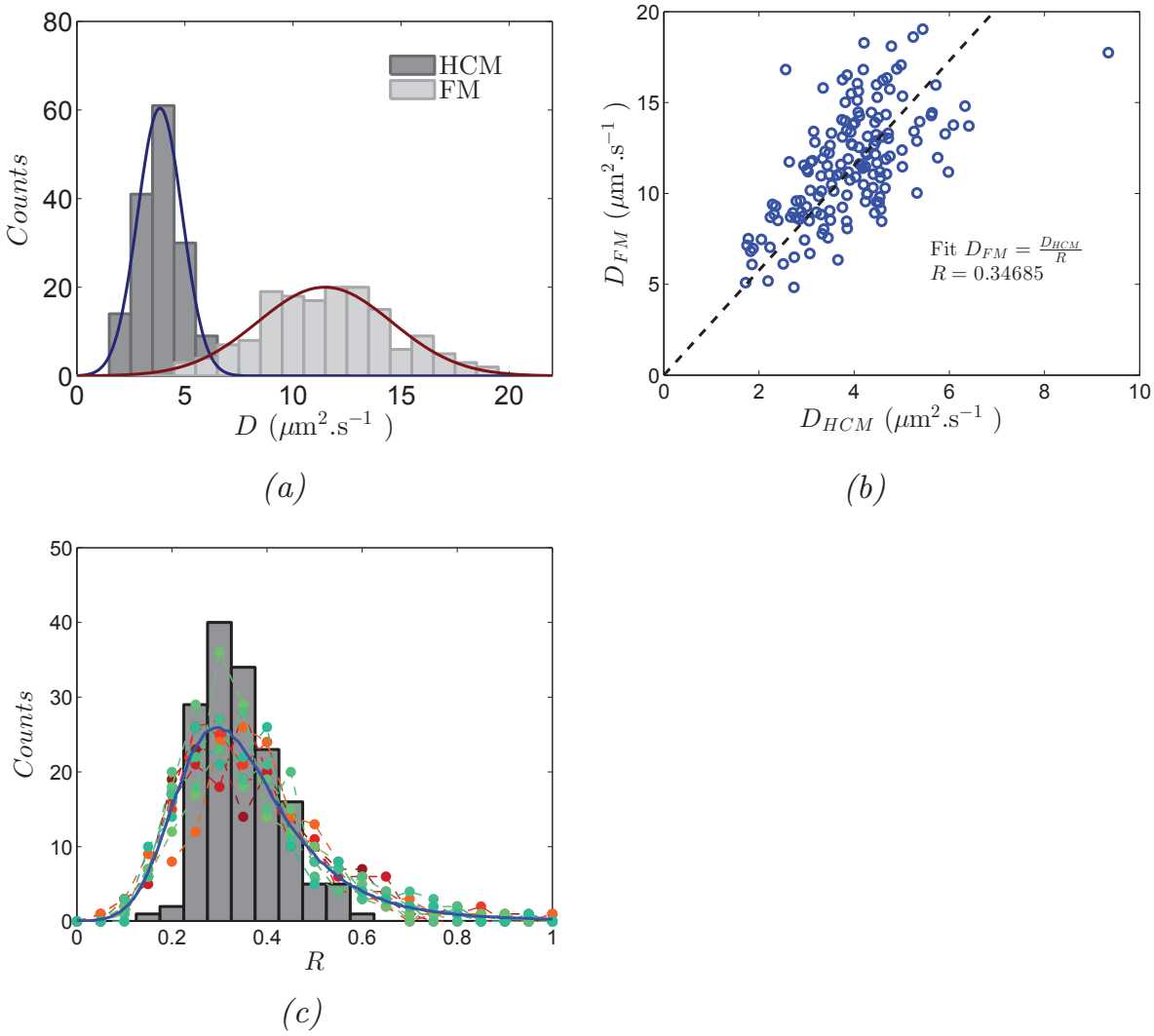


Figure 3.21: Overview of the diffusion constants in the dataset. (a) Distribution of D in FM (red) and HCM (blue). The solid lines are Gaussian fits. (b) A linear relation between D_{FM} and D_{HCM} looks realistic, validating the use of the relative diffusion constant R as a measure of the QSF response. (c) Distribution of R (bars) and prediction if D_{FM} and D_{HCM} were uncorrelated (blue line). The red (resp. green) disks show six realisations of the distribution of 156 computed R from simulated D values (resp. experimental, decoupled D values).

During these experiments, that have been done between May and December 2015, we could observe a certain variability in what we call the "basal" cell motility, namely the value of D in FM. It leads to notably spread values of D both in FM and HCM (see 3.21), which form almost a continuum from 2 to 20 $\mu m^2.min^{-1}$. Their respective distribution are correctly fitted by normal laws with $D_{FM} = 11 \pm 3\ \mu m^2.min$ and $D_{HCM} = 4 \pm 1\ \mu m^2.min$ respectively, where the figures before and after

the '±' are the mean and the standard deviation. Yet, D_{FM} and D_{HCM} are quite correlated. Indeed, we find a correlation value of $Corr = 0.6$, where the correlation is defined as

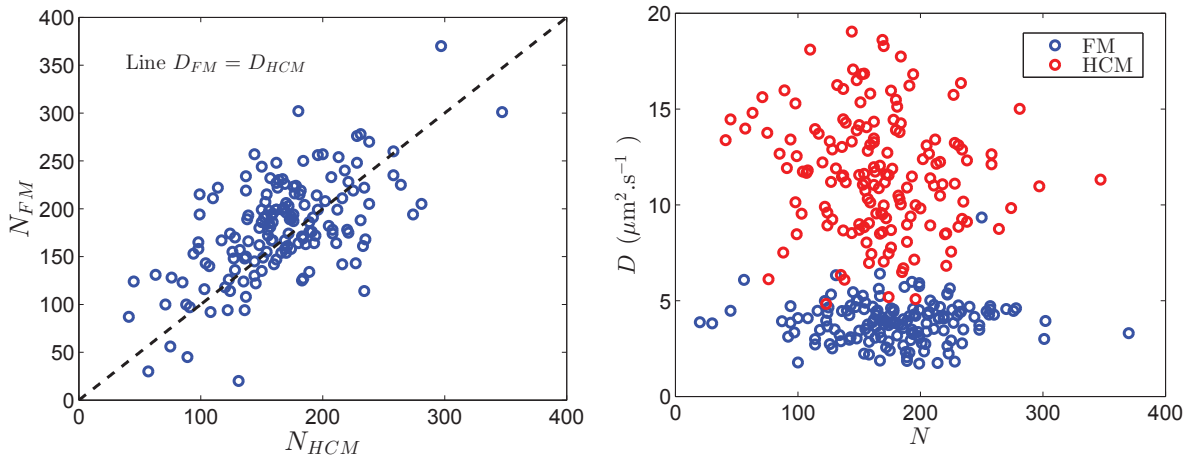
$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X).Var(Y)}}. \tag{3.16}$$

We computed the relative diffusion constant R and found that it is distributed according to Fig. 3.21. In particular, its mean and standard deviation yield $R = 0.34 \pm 0.08$. We checked further this correlation by contrapositive. To that end, we assumed that D_{FM} and D_{HCM} were distributed normally with their measured mean and standard deviation, but uncorrelated. In that case, the distribution of R should be a complicated function (see [94]). Instead of computing directly this function, we generated 10^6 uncorrelated couples of random values following the distributions of D_{FM} and D_{HCM}, then computed their ratio. The obtained R_{sim} variable is distributed around a similar average as the experimental one, $\langle R_{sim} \rangle = 0.37$, but the tails are broader (Fig. 3.21c) due to the lack of correlations. This is not a low-number artefact, as shown by the distribution of R computed from the experimental values, and decorrelated by muddling up randomly the values of D_{HCM}, which is fairly similar to the theoretic, uncorrelated one (Fig. 3.21c).

This is a first illustration of the *exogenous* variability that differs from the *endogenous*, QSF-caused, effects.

Density is not a control parameter. Of course, other exogenous parameters could play a role, and it is important to rule out the main observable one: the cell number. Despite our efforts to keep the cell density constant, the actual cell number varied from less than 100 to more than 300. Interestingly, the relation $N_{FM} \sim N_{HCM}$ is quite well verified (Fig. 3.22a), which reduces the variability on the $\frac{N_{FM}}{N_{HCM}}$ ratio, hence on the intra-experimental variability caused by this parameter, if ever. Moreover, it is quite clear from the D *vs* N plot (3.22b) that in these very dilute situations, the cell density has a negligible effect on the motility.

The speed is neither a constant, nor the only affected parameter. Then we checked whether the hypothesis of constant speed held in our experiments. The answer is clearly no: D appears to be a function of v, and one could even reasonably assume a linear relation between v^2



(a) *(b)*

Figure 3.22: D is not correlated to the cell number in dilute conditions. (a) N_{FM} versus N_{HCM}: the variations seem to be more inter-experimental than intra-experimental. (b) D is not correlated with N.

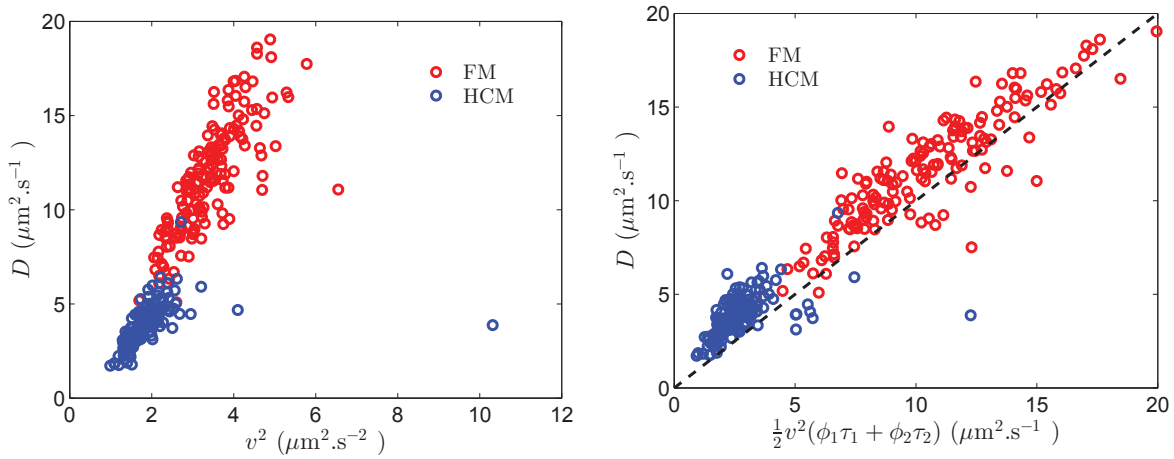
and D. Remember that in the simple CRW model in 2D, such a relation exists:

$$D = \frac{v^2 \tau}{2}, \tag{3.17}$$

where τ is the persistence time. Should this kind of relation hold in our more complex case, and should the QSF act only on the cell speed, then we should get $R = \frac{D_{HCM}}{D_{FM}} = \frac{v^2_{HCM}}{v^2_{FM}}$. In other words, the relative diffusion constant should be accounted by the relative squared speed $R^2_v = \frac{v^2_{HCM}}{v^2_{FM}}$. It is not the case and R is only loosely correlated to R^2_v – with $Corr(R, R^2_v) = 0.2$ (scatter not shown). Furthermore, the relative squared speed R^2_v is higher than the relative diffusion constant: $R^2_v = 0.56 \pm 0.20 > R = 0.34 \pm 0.08$. It means that a part of the regulation of D by the QSF is not accounted for by its action on the cell speed. According to 3.17 or its suitable analogue, the persistence of the movement should also be affected, as first claimed by Golé *et al.* [18].

Velocity autocorrelations. To study the evolution of persistence in these experiments, we computed the velocity auto-correlation function, defined as:

$$C(\delta t) = \langle \frac{\mathbf{v}(t+\delta t).\mathbf{v}(t)}{\|\mathbf{v}(t+\delta t)\|.\|\mathbf{v}(t)\|} \rangle. \tag{3.18}$$



(a) *(b)*

Figure 3.23: (a) D is quite correlated to v^2 both in FM (red) and in HCM (blue). (b) D is well predicted by equation 3.20.

As detailed in **Chapter 2** of the present manuscript, in the case of simple correlated random walk, this function decays exponentially with a characteristic time τ called the persistence time, which is the τ used in relation 3.17. It corresponds to the time over which, in average, the cell loses memory of its previous direction of motion. Our data are best fitted with a sum of two exponentials:

$$C(\delta t) = \phi_1 \exp(-\frac{\delta t}{\tau_1}) + \phi_2 \exp(-\frac{\delta t}{\tau_2}). \tag{3.19}$$

Because $\phi_1 + \phi_2 \neq 1$, this highlights the existence of *three* different relaxation times, and hence three modes of migration coexisting in the population: a fast relaxation with persistence time lower than the frame rate $\delta t = 20$ s which can be at least partly due to pixelation or positional noise, an intermediate persistence τ_1 and a larger persistence τ_2. Their proportions in the population are $\phi_0 = 1-\phi_1-\phi_2$, ϕ_1 and ϕ_2 respectively.

Given the number of parameters, it seems more reasonable to set two of them to fix values in order to obtain valuable insight from the two free ones. We found that keeping $\tau_1 \in [1;3]$ min and $\tau_2 \geq 7.5$ min yields both a good fit of the data and reproducible behaviours for ϕ_1 and ϕ_2.

Up to some assumptions – exponential distributions for the times spent in each respective mode is a sufficient condition for instance [77] – the analogue to equation 3.17 in this three-modal situation writes:

$$D = \frac{v^2(\phi_0\tau_0 + \phi_1\tau_1 + \phi_2\tau_2)}{2}$$
$$\approx \frac{v^2(\phi_1\tau_1 + \phi_2\tau_2)}{2}, \text{ if } \tau_0 \simeq 0. \quad (3.20)$$

Plotting D *vs* the approximated right-hand-side of 3.20 shows the remarkably good accuracy of that prediction (Fig. 3.23b). To compare, the values of D computed with a single relaxation time (but still a $\phi < 1$) are quite under-estimated, even though they correlate well to the actual measurements.

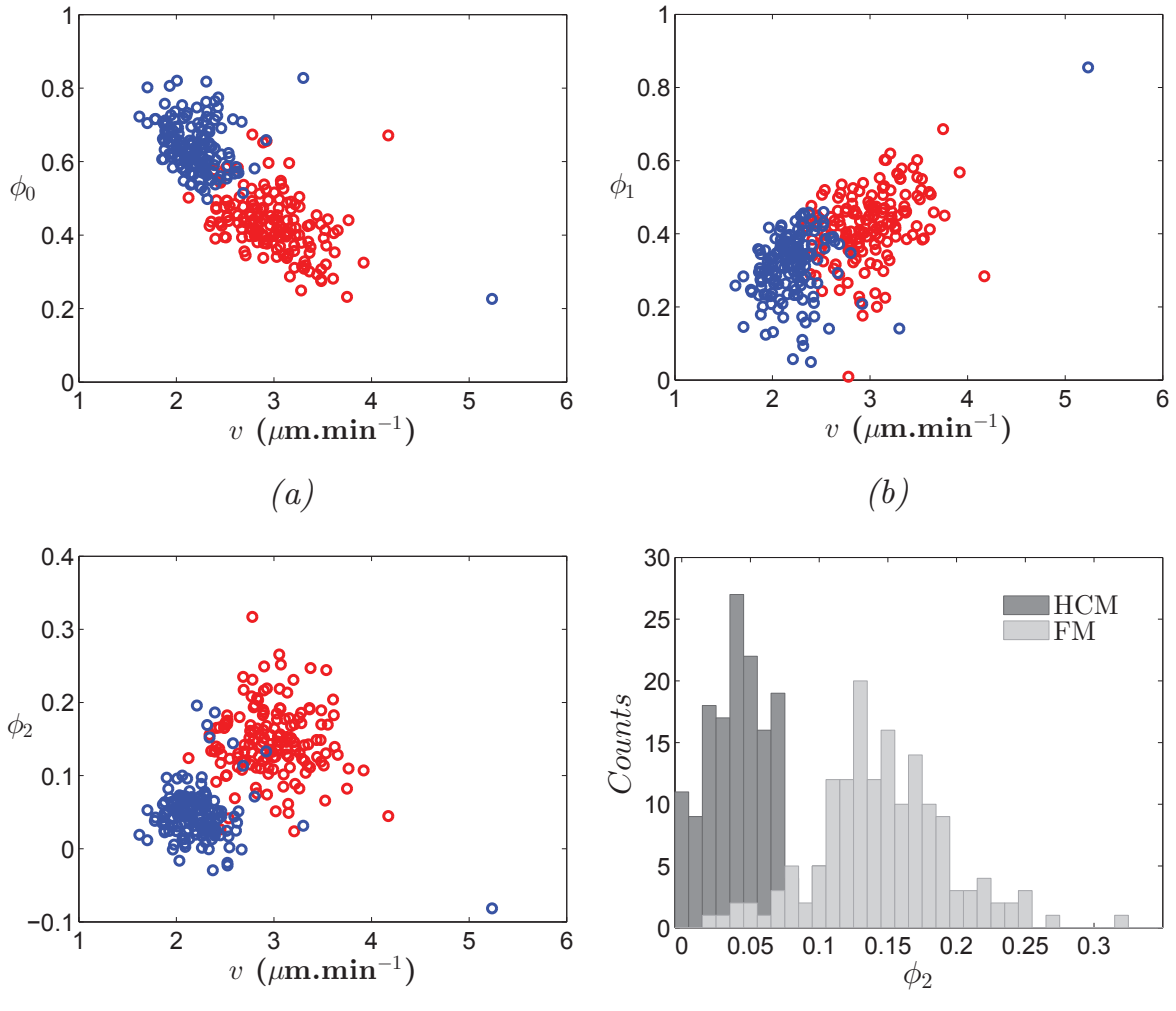


Figure 3.24: Regulation of the persistence modes ϕ_i and relation to v. (a) ϕ_0 exhibits a strong negative correlation to v. (b) ϕ_1 is positively correlated to v. (c) The ϕ_2 values in FM and HCM are well distinct. However, in each of the clouds, there seems to be no correlation of ϕ_2 with v. (d) Distribution of ϕ_2 in FM (light gray) and HCM (dark gray).

The three 'ϕs' are correlated with D and v when all the data are considered (see Table 3.1 and Fig. 3.24). These correlations are reduced when the data in FM and in HCM are separated, suggesting that a part of the correlation is due to QSF-related variations of all the parameters. Strikingly, though, the correlations are even more strongly destroyed in the case of ϕ_2 (Table 3.1 and Fig. 3.24). It could mean that ϕ_2 is regulated by the QSF in a way that is independent of the other parameters, and especially the cell speed. It is anyway strongly reduced (three times on average, Fig. 3.24d), confirming that the QSF acts on the cells' persistence, by decreasing the occurrence of long-lived polarised motion.

Corr	D	v	D (FM)	D (HCM)	v (FM)	v (HCM)
ϕ_0	-0.83	-0.76	-0.52	-0.45	-0.37	-0.43
ϕ_1	0.55	0.58	0.30	0.34	0.32	0.41
ϕ_2	0.72	0.53	0.27	0.06	-0.03	-0.17

Table 3.1: Correlation coefficients of the distribution of the ϕ parameters with D and v, either gathering all the data (first two columns) or separating FM and HCM data (left columns). The figures in red (resp. blue) denote significant positive (resp. negative) correlation,

6 Conclusions and perspectives

A few years ago, Laurent Golé discovered the existence of a secreted QSF that regulated the motility in DH1 cells. Here we extended this finding to the AX2 cell line, with the ambition of understanding the chemical nature and the mechanism of action of the QSF. We found out that this QSF system also exists in AX2, with strikingly conserved properties: even though the basal – "natural" – motility of AX2 cells is relatively different from that of DH1 cells, and especially in so that the cells move faster, the response to the QSF in terms of effective diffusion constant has the exact same dependency on the QSF concentration.

We could unravel the molecular functioning of this system: it is based on a secreted amphiphilic and uncharged molecule of high molecular weight, namely the QSF; this QSF induces a response *via* G-protein subunits, particularly Gα4 but also Gβ1; for now, the downstream signalling pathway remains puzzling. This response manifests itself as a decrease in both cells' speed and persistence. Regarding the latter, the regulation is especially visible in an average 3-fold decrease of ϕ_2, the proportion of 10 min persistence in the trajectories. This regulatory system is apparently linked to non-linear intracellular dynamics. Indeed, the motility decreases exponentially with the concentration towards a low plateau.

At the other end of the chain, the secretion itself undergoes a feedback that leads to a linear decrease of the secretion rate with the QSF concentration. This results in a surprising behaviour: the QSF concentration c transitions very fast from a low value to its saturation, in a way that is essentially independent from the past history of the colony, but rather induces an almost bijective relation between c and the instantaneous cell

density. It raises the question of possible other functions of the QSF: as such, it would seemingly constitute a very robust self-counting system, with a narrow detection range around a density value that is known to be a watershed in *Dictyostelium* culture.

In a nearer future, other experiments will be necessary to complement the already done work. This is especially true for the uncovering of the molecular response pathway, which might require a larger scale screening to find non-responding mutants. The complete biochemical identification of the QSF could also take time: a battery of new, more focused tests, *e.g.* treatment with periodate which is known to degrade most carbohydrates, has to be implemented to that aim. In the meanwhile, there is still room for more physics-based work, particularly on the effect of the QSF on the protrusion dynamics.

---

4. Spreading of model colonies

This chapter presents an analysis of the spreading of model cell colonies. To provide an experimental set-up that mimics various complex physiological situations, such as the tumour escape, the ecological spreading or even some morphogenetic movements, in a very minimal way, we used a micro-stencil technique to initially locate *Dictyostelium* cells at a desired density in a disk with controlled diameter. Then, we released this constraint to let the cells move freely outwards. We relate the macroscopic evolution of the colony to the properties of individually tracked cells. In particular, we dissect the respective effects of proliferation, cell motility and various kinds of cell-cell interactions. We find that proliferation is dominant in the long-term behaviour, together with a probable noticeable overall effect of the QSF which is described in details in the previous chapter. At short-time, by contrast, short-range interactions induce a density-dependent speed-up of the spreading. More precisely, cell-cell contacts seem to increase the persistence of the cell trajectories, a previously unaccounted effect that we term *Contact Enhancement of Locomotion* (CEL). Our analysis of the experimental data is supported by individual-based models and mean-field partial differential equations descriptions of the system.

Contents

1 Introduction

The collective regulation of the motility is particularly relevant when it comes to the following question: how will cells in a place move to another place? Especially, will the population move as a whole or by sending isolated scouts ? Will the dynamics of the process be influenced by the properties of the initial group? Such issues are at play in various biological situations, inspiring either ecologists (how will some cell population invade its surroundings?), developmental biologists (why should this group of cells be translocated to that particular place in such a way?) or oncologists (on what grounds is one tumour more or less invasive?). To determine the underlying principles of such complex processes, physics suggests to first try to understand the basic ingredients controlling a minimal caricature of them. The study of the spreading of patterned circular colonies of cells seems to fit well this recommendation, all the more with cells that do not adhere specifically onto each other, which would add an inopportune particularism.

In the following, we review the current knowledge about cell colony spreading, in the light of previously realised experimental observations. We stress out the accepted relative importance of cell proliferation and motility. Then we focus on already described mechanisms of cell-cell contact interactions, as we will see they can play a prevailing role. The next sections present our experimental procedure, the analysis of our observations, and some insight provided by several modelling approaches.

1.1 Cell colonies: from bacteria to tumor

In the introductory chapter of this thesis, we showed how the cell proliferation can be regulated, leading to changing dynamics. The temporal dynamics being quite clear, let us now introduce space in a simple fashion. Early papers on colony spreading, dating back from the 1960s [95, 96, 97] dealt with the spreading of microbial – mostly bacterial – colonies. These works did not really consider the matter of motility, but rather focused on growth-governed and nutrient-diffusion-limited patterns.

Although very simple in their principle, these models based only on proliferation introduce two fundamental features that we always encounter, even in more complicated situations. These are the concepts doubling time and peripheral “growing zone”, which almost direct consequences are an exponential or linear growth of the colony radius. To visualize it, let us consider the mass increase of a circular 2D colony, or equivalently

of a cylinder of constant height h. In that view, the mass transport is implicit and may be forgotten, and the radius of the colony is simply related to its mass:

$$M = \pi R^2 h \Leftrightarrow R = \sqrt{\frac{M}{\pi h}}. \tag{4.1}$$

One can then study two simple situations (Fig. 4.1). First, all the cells proliferate. The dynamics is given by the differential equation 4.2, yielding an exponential increase of the radius, of which the time constant is related to the cells' doubling time 4.3:

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \lambda M \Rightarrow M = M_0 e^{\lambda t} \tag{4.2}$$

$$\Rightarrow R = R_0 e^{\frac{\lambda t}{2}}, \tag{4.3}$$

where M_0 and R_0 are respectively the initial mass and radius, and λ the growth rate.

In the second case, suppose that for any reason, only a ring of width W at the edge is able to proliferate. That can be due to the consumption of nutrients in the core of the colony, or to a saturation of density in the central region for instance. Then the dynamics for the mass is given by the following equation, leading to a linear increase in the colony radius [95] (see Fig.4.1):

$$\begin{aligned} \frac{\mathrm{d}M}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\pi R^2.h) = 2\pi h R \frac{\mathrm{d}R}{\mathrm{d}t} = \lambda 2\pi W h \\ \Rightarrow R(t) = R_0 + W\lambda t. \end{aligned} \tag{4.4}$$

We will see later (see next section) that these features are recovered when motility is explicitly considered, and especially that the notion of peripheral growing zone can be explain more precisely in that scope.

The interest in the spreading of colonies of cells has risen together with the growing concerns about cancer. Interestingly enough, some features of these systems are reminiscent of the periphery versus core discrimination. For instance, it is known that in non-vascularised tumours, a necrotic core forms and that the spheroid properties are much affected by this duality between the inside and the active envelope.

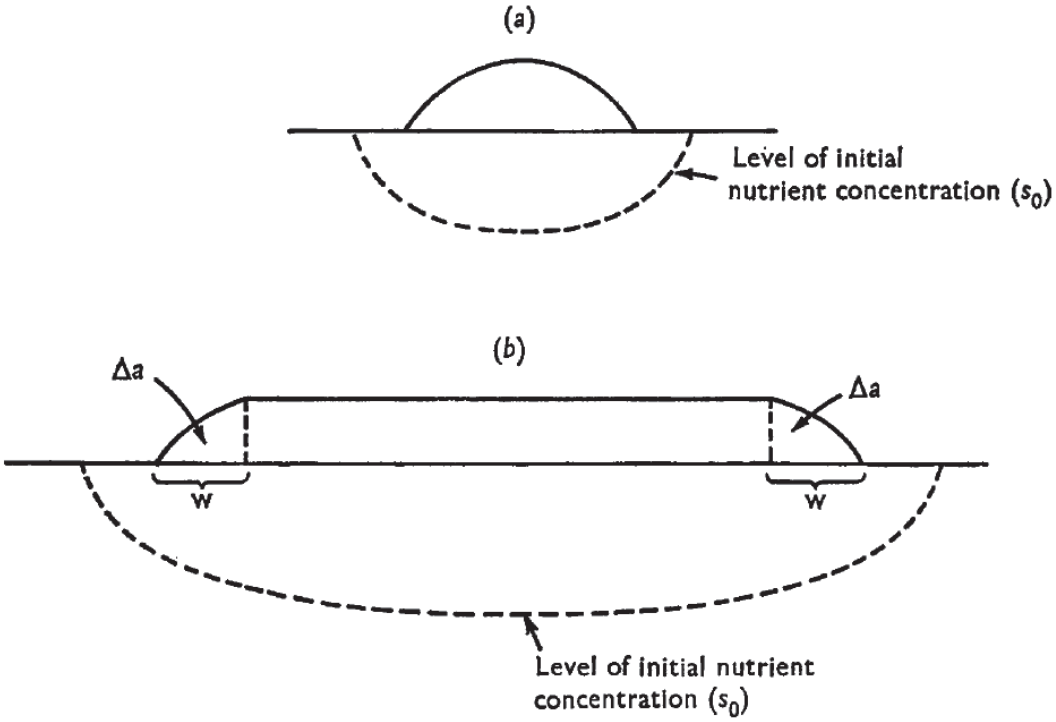


Figure 4.1: Sketch of the bacterial growth model from [95]: first all the cells proliferate (top), then only a ring of width W is able to grow. here, the nutrient concentration is considered as the limiting factor for growth.

Subsequently, huge modelling efforts have been made from the 1980s [98]. From that point, the cell motility has started being taken explicitly into account in the theoretical descriptions of expanding cell populations.

1.2 The decadence of a too powerful model, or trendy Fisher waves

The FKPP equation. Knowing the properties of basic ingredients that can affect the evolution of cell density, namely proliferation and motility, the next step is to try to write a mean-field model for this density. The usual way to derive such an equation is to start from a simple mass conservation equation (where ρ represents the cell density, $\mathbf{j}$ the cell current and λ a ρ-dependent reproduction rate).

$$\frac{\partial \rho}{\partial t} = -\boldsymbol{\nabla} \cdot \mathbf{j} + \lambda(\rho)\rho. \tag{4.5}$$

Then one needs to write the flux and production terms in terms of the density to close the equation. For the flux, the simplest way to integrate a correlated random walk in a field equation is to use the first Fick law

– $\boldsymbol{\nabla} \cdot \mathbf{j} = -D\Delta\rho$ – and hope the persistent aspect is not to important. Concerning the production, a logistic term appears to be the simplest form that can account for the known growth dynamics (although one would prefer keeping things linear, a simple linear proliferation would produce an unrealistic diverging density). We arrive then at the following model:

$$\frac{\partial \rho}{\partial t} = D\Delta\rho + \lambda\rho\left(1 - \frac{\rho}{\rho_0}\right). \tag{4.6}$$

This equation was first introduced by Fisher [99] to describe the spread of a favoured gene in a population, and analytically studied by Kolmogorov *et al.* [100]. It is known as the Fisher-Kolmogorov-Petrovskii-Pishkunov (FKPP) equation and it has been quite deeply reviewed in Murray's book [101].

At low densities, it is similar to a simple diffusion equation, with uniform exponential growth, since the proliferation term is then simply linear. Thus, the spreading of a low density colony should obey usual diffusion laws, like the time square-root increase of the gyration radius. Depending on the parameters, the density will either increase from the beginning or first decrease before eventually increasing again. Indeed, starting with N_0 cells in a domain of area A_0, the area scales as:

$$A \sim A_0 + 4Dt, \tag{4.7}$$

and, provided $N_0/A_0 \ll \rho_m$, the cell number scales as

$$N \sim N_0 e^{\lambda t}, \tag{4.8}$$

and so the cell density is given by

$$\rho = \frac{N}{A} \sim \frac{N_0 e^{\lambda t}}{A_0 + 4Dt}. \tag{4.9}$$

Differentiating with respect to time, one finds that

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} \geqslant 0 \Leftrightarrow \lambda A_0 \geqslant 4D. \tag{4.10}$$

If the colony spread fast enough, it can offer enough new area per unit time for the newly produced cells, so that the density first decreases; otherwise it increases from the beginning.

Fisher waves. Then the density approach the carrying capacity. The behaviour is somehow different, but the density still increases until it saturates, and the population still spreads out due to the negative gradient. When the density reaches the carrying capacity, starting from the center, the situation is the following: in the central zone the density profile is flat, equal to the carrying capacity, and so it remains constant (no proliferation, no gradient); outside this zone, there is a proliferating, spreading zone (Fig. 4.2). This situation is very reminiscent of the peripheral growing rim assumption of the bacterial models reviewed in the previous section [95].

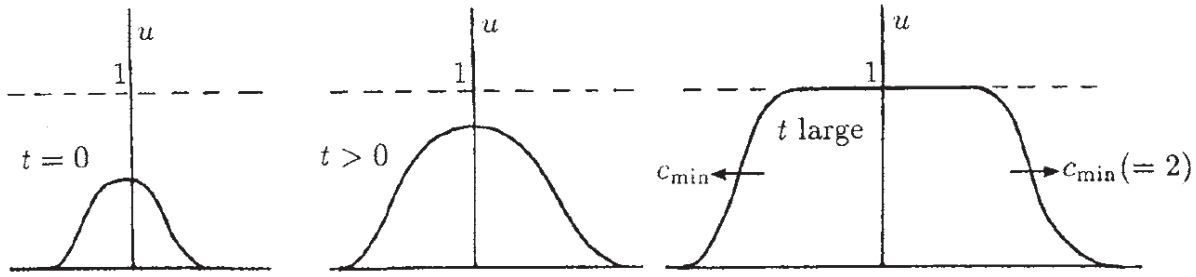


Figure 4.2: Dynamics of the adimensioned Fisher-Kolomogorov-Petrovskii-Pishkunov equation in one dimension (from [101]). Starting from a "spot" (left), the density u spreads out and its integral grows (center), until it reaches the carrying capacity $u = 1$ and a wavefront develops to invade the empty space at speed c_{min} (right).

It has been shown [101] that in 1D the FKPP-equation admits travelling wave solutions. Several wave velocities are accessible but the slowest one, given by $c = 2\sqrt{D\lambda}$, is kinetically selected. This wave, whose shape depends on the parameters, separates a region where the density is equal to the carrying capacity from another where it is zero, the former invading the latter. One can guess that this behaviour can map to the higher dimension radial symmetry situation for large enough radii, giving rise to a linear increase of the colony radius. Actually, this latter result is true even for non-negligible curvatures (one can see that from numeric resolution for instance), and the front wave tends asymptotically to that of the 1D case (see [98] p444).

Numerical simulations. Therefore, if this equation holds, one should notice a transition from $R \sim t^{1/2}$ to $R \sim t$, where R denotes the colony radius. This transition has been observed in numerical simulations of randomly moving particles. Dirk Drasdo and his collaborators has produced a substantial literature [102, 103] about this kind of individual

based models. Considering hard or soft, disk-like, randomly moving cells, they obtain results that are quite consistent with the Fisher model. In particular, they question the size of the proliferating ring with more individual-based considerations.

In the case of adherent cells, used to model connective monolayers, they also point out a transition from an initial exponential increase of the radius to the linear long-time behaviour. This exponential evolution, instead of square-root, is due to the suppression of free diffusion by the cell-cell adhesion. Interestingly, this initial exponential spreading has been recently observed experimentally [44], with some subtle modifications, for circular colonies of MDCK cells, using a set-up that is very similar to the one presented in this Chapter.

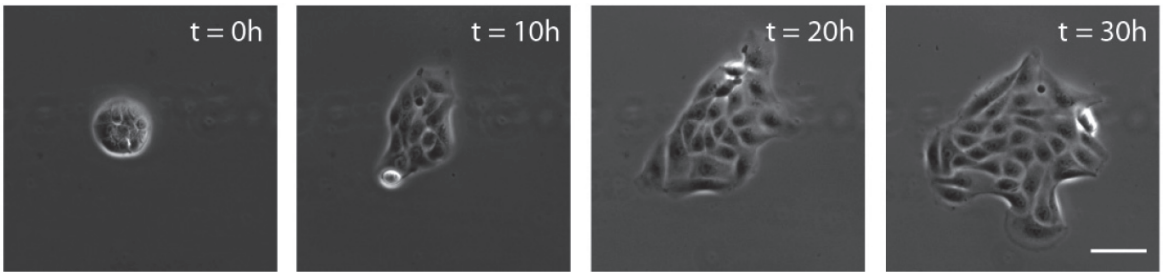
Experimental works. In the last two decades, in addition to the work of Marel *et al.* [44] on islets of MDCK cells (Fig. 4.3a), a few authors [104, 105, 106, 107, 108, 109, 110, 43] experimentally studied the spreading of circular colonies of crawling cells, using initial radii that are notably larger than the ones we use ourselves (Fig. 4.3b). All of them observe approximately a long-time linear increase of the colony radius, as predicted by the FKPP equation.

Some of them [104, 105, 106, 107, 108] then went to an analysis of the critical exponents of the interface growth to try to classify their system in a known universality class of growing interface, which is quite far from our present concerns.

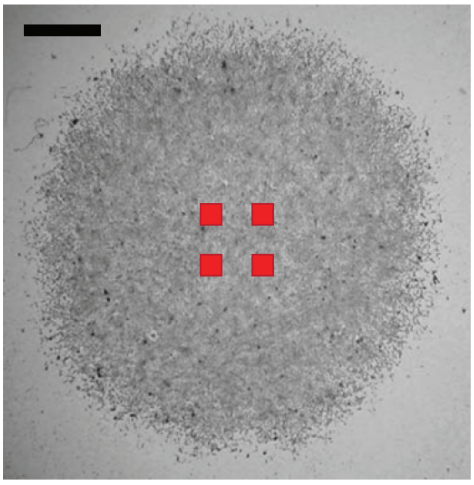
Lange *et al.*, as Simpson *et al.* [43, 110] just took this model for good, and tried to fit in without examining too much its validity. The latter, for instance, measured λ and c, the linear spreading rate, to deduce the cell diffusion constant from the formula $c = \sqrt{4\lambda D}$. Yet they did not look for any microscopic confirmation of the obtained value. On the contrary, Lange *et al.* tracked the single cells and they even quantified some deviations from the assumptions of the FKPP equation: the cells on the border are more persistent and preferentially directed outwards than the cells inside. Their point being not to question this model, though, they do not stress that out.

Sengers *et al.* [109] had an opportunity to go further, as they had to modify the flux term in what they call the "sharp front version" of the FKPP equation,

$$\frac{\partial \rho}{\partial t} = \boldsymbol{\nabla}.(D\frac{\rho}{\rho_{max}}\boldsymbol{\nabla}\rho) + \lambda.\rho.(1 - \frac{\rho}{\rho_{max}}), \tag{4.11}$$



(a)



(b)

Figure 4.3: Examples of previous cell spots experiments.
(a) Islet of MDCK epithelial cells (from [44]). Scale bar 100 µm.
(b) Large colony of murine 3T3 fibroblasts (from [110]). Scale bar 1.5 mm.

to fit correctly their experimental density profiles. They claim two possible origins for this flux expression: random motion with increased speed in higher density areas, or directed motion proportional to the "gradient of free space" $\boldsymbol{\nabla}(1 - \frac{\rho}{\rho_m})$, but they do not support either hypotheses by cell motility data.

Lastly, Marel *et al.* [111] also used the Fisher model with a modified flux expression, in a channel geometry. There, they added a constant drift flux in Fick's first law to account for their flux data whose dependence on the density gradient is very small. Then, they deduced a constant drift *velocity*, which is very different in truth, to define a moving frame where they just fitted their (almost linear) density profiles with an analytic solution of the FKPP equation (actually, with the almost linear part of a complicated deformed-sigmoid) to obtain the cells' division time and diffusion constant, which, nevertheless, are consistent with microscopic measurements.

As a conclusion, the Fisher model is the simplest mean-field equation that one can think of for the density of randomly moving, logistically growing cells. It has been studied for long and some interesting results

have been obtained on it. While some of its features are very satisfactorily found in individual-based models, as well as in some experimental observations, it has not really been subjected to experimental assessment in the field of cell colony spreading, even though some papers [109, 111], willingly or not, started to point out some contradictions with it, clearing the start of possible tracks beyond the FKPP paradigm.

1.3 Scope of our experiments

In the previous sections and in **Chapter 1**, we gave an overview of the two main processes at stake in the spreading of colonies of motile cells. We showed that there are solid foundations and a scientific consensus about the description of cell proliferation with a logistic law, characterized by a doubling time and a carrying capacity; similarly, single cell motion can be described in a quite satisfactory fashion with a persistent random walk characterized by its speed and persistence time.

Some models, reviewed above in this section, proposed descriptions of cell population dynamics that takes these mechanisms into account. Among them, the Fisher model, or FKPP equation is at the same time simple and quite complete, which granted it a long-lived success. Although some experimental works are consistent with it to some extent, they still lack a microscopic assessment to confirm its validity in a quantitative way.

In particular, it considers no interactions in the motility term. Yet, it is clear, both from an experimental and a theoretical point of view (see **Chapter 1**), that cells interact together and that these interactions should affect the local motility properties, such their speed or polarisation.

To check the role of these interactions and the agreement of FKPP-like models to a comprehensive, coherent, set of experimental data, we designed experiments that allowed us to access the properties of motion and proliferation in spreading colonies of cells, over a wide range of spatial and temporal scales, with a good control on parameters such as the cell density, the population geometry and the environmental conditions.

In what follows, we first describe the experimental set-up and analysis framework. Then the experimental outputs are presented in details. Last, before we conclude on that part, we introduce some modelling approaches, in which view we discuss our results.

2 Methods

2.1 Design of the experiment: prerequisites

The purpose of the experiment is to decipher the roles and regulation of motility and proliferation in a colony assay. To that end, we decided to fabricate colonies by the suited patterning technique. We used a stencil technique that seemed to be the most efficiently implementable at many scales. This allows us to control the initial shape and dimension, as well as the cell density, of the experimental colonies.

The radial symmetry was chosen as (i) this is the most stable symmetry for spreading finite-size systems and it occurs naturally in spreading colonies [112], (ii) it allows easy conceptual analogies to diffusion for instance (iii) it allows orthoradial averaging provided the symmetry is not broken.

The diameter of the colonies was first set to be tunable from 40 µm to 320 µm, allowing to investigate several regimes from two-body to true multicellular – even though still mesoscopic – assembly, but only 320 µm was eventually used. This size is large enough to consider a wide range of cell densities, but small enough (i) for the cell number to remain reasonable, (ii) to allow the colony to double in area quite fast with cell motion only, (iii) to explore the transition between initially decreasing and increasing density due to the interplay of spot spreading and cell number increase.

We decided to make only one colony per experimental well to ensure the absence of external cues. To that end it was also important to make sure that no cell (or just a few of them) settled outside the colony, and that the medium volume was large enough for the colony not to feel the sample walls.

To allow a better control of the chemical composition of the medium, we should also be able to infuse the sample through a fluidic system.

We chose to fabricate micro-stencils made in PDMS after the technique used by Poujade *et al.* [113]. Here we describe the sample preparation, from mould fabrication to cell deposition, and the bases of time-lapse data analysis. The development of the sample preparation protocol was made with the collaboration of Peiyuan He, a Master 1 intern under my supervision, from January to February 2014.

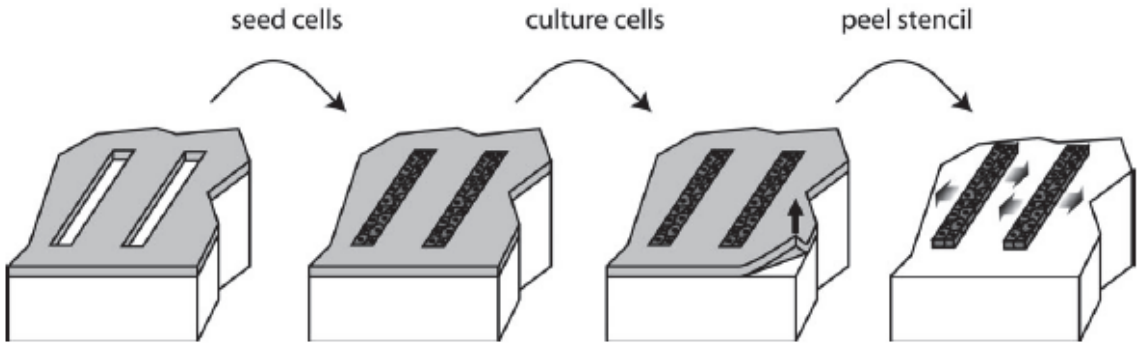


Figure 4.4: Principle of the stencil technique: the cells are first seeded on top of a PDMS micro-stencil, then, after they could adhere to the substrate through the stencil's holes, the constraint is removed to let the cells move freely (reproduced from [113]).

2.2 Soft lithography

To make PDMS stencils, we needed a mould. We used standard soft-lithography processes to fabricate a SU8 mould containing an array of pillars, whose negative PDMS output would be a membrane with regularly spaced circular holes. The SU8 needed to be thick enough to cross the entire PDMS layer.

The 2-steps process described here is for an expected height of ~ 500 µm for the SU8 pillars (Fig. 4.5). We used also a $160-170$ µm-thick mould that was made following the same process with only one spin-coating step.

2.3 Micro-fabrication: from PDMS scrapping to PDMS spin-coating

The crucial point is to be able to make transverse holes in the PDMS. So the PDMS layer must be thinner than the SU8 features, which in turn cannot be more than 500 µm-high as we saw.

We first tried a scrapping device that allows pouring PDMS into a layer of controlled height. Although imprecise and uneasy to use, this device was used to make the first stencils. Two different problems occurred at that point. First, some of the pillars were torn off while peeling off the stencils from the mould, likely due to an inverse conical shape or just because they were tilted from the verticality, leading to an obvious non-durability of the mould, and a difficult tweezers extraction of the pillars that remained stuck in the PDMS membranes (Fig. 4.6a). Second, at the cell seeding step, the cell suspension had a hard time getting down

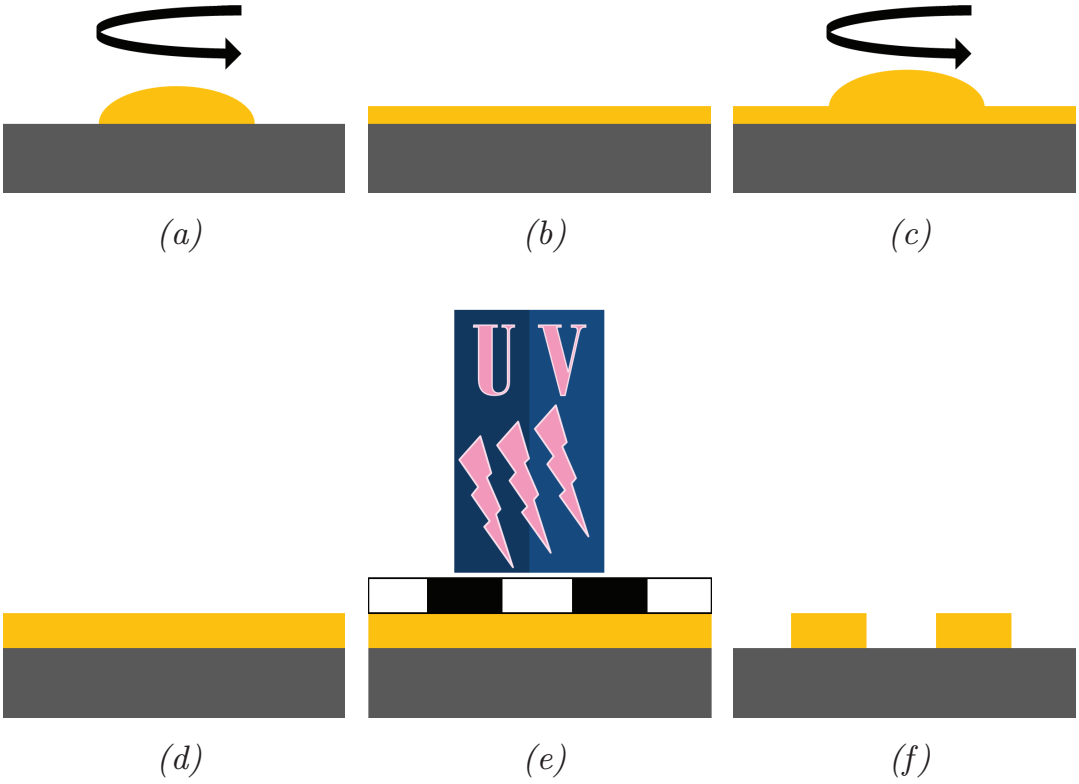


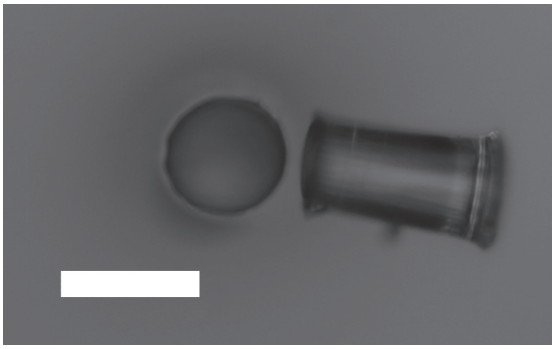
Figure 4.5: Photo-lithography.
(a) Photoresist is deposited and spin-coated on a Si wafer. (b) Then it is cured by heating. (c) A second layer of photoresist is deposited and spin-coated on top of the previous one. (d) The second layer is cured. (e) The photoresist is subjected to UV illumination through a photo-mask. (f) The patterns are developed in a chemical bath after a postbake.

the 500 µm-high but only 320 µm-wide chimney through the PDMS.

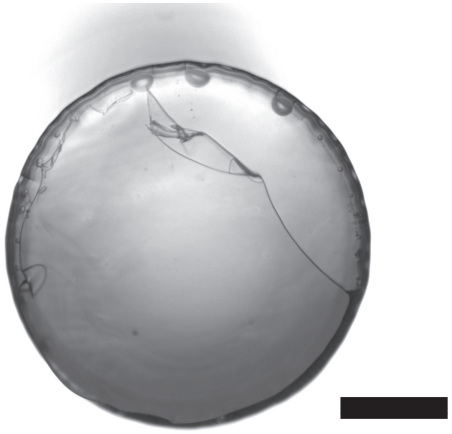
These two considerations together lead us to think about making things thinner, hence the idea of PDMS spin-coating. This technique had already been used and documented [114, 115], but its drawback in the present case is the possibility of wake formation behind the mould's features, leading to non-uniformity of the surface. The feature density being low on our design, this risk was not too high, though.

We used a 200 µm-high mould for this, with lower aspect ratio, hence more robust pillars. PDMS (10:1 base:curing agent ratio) was then spin-coated on this wafer (500 rpm, 10 s, then 800 rpm, 1 min). It resulted in a 70 µm-thick layer of PDMS pierced by the SU8 pillars (Fig. 4.6c).

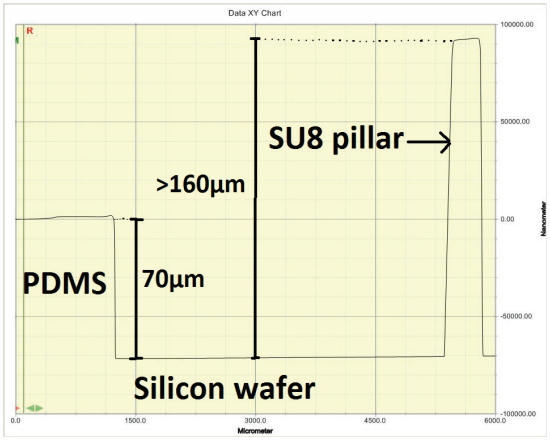
After peeling of the PDMS membrane in the shape of individual squares of 9x9 mm with one single 320 µm-wide hole in the center, a very thin membrane of PDMS usually remained on top of the hole. We removed it with the tip of a scalpel under a bright-field microscope at 10X magnification (Fig. 4.6b).



(a)



(b)



(c)

Figure 4.6: PDMS stencils. (a) Top view of a PDMS stencil with a 80 µm*-wide hole and a pillar pulled up from the mould. Scale bar* 100 µm*. (b) Top view of a PDMS stencil with a* 320 µm*-wide hole. The thin PDMS membrane that obstructs the hole is visible. Scale bar* 100 µm*. (c) Profilemeter image of the mould (side view). Left: spin-coated PDMS. Central part: free mould surface from which the PDMS has been peeled off. Right: SU8* 320 µm*-wide pillar. (Warning: the scales on the x-axis and y-axis are different!)*

The resulting PDMS squares were stored in plastic dishes closed with Parafilm M (Pechiney plastic packaging, Chicago, IL) to protect them from the dust and moisture before use.

2.4 Sample preparation: plasma or not plasma, glass or plastic, these are the questions. Bonus: the benefits of vacuum

The exact nature of the sample dish is discussed just below. Let us picture it as a circular, few-centimeter-sized, transparent experimental dish. A PDMS square was stuck in the center of each experimental dish using its natural stickiness, probably helped by electrostatic interactions. To avoid the deposition of a large amount of cells around the PDMS

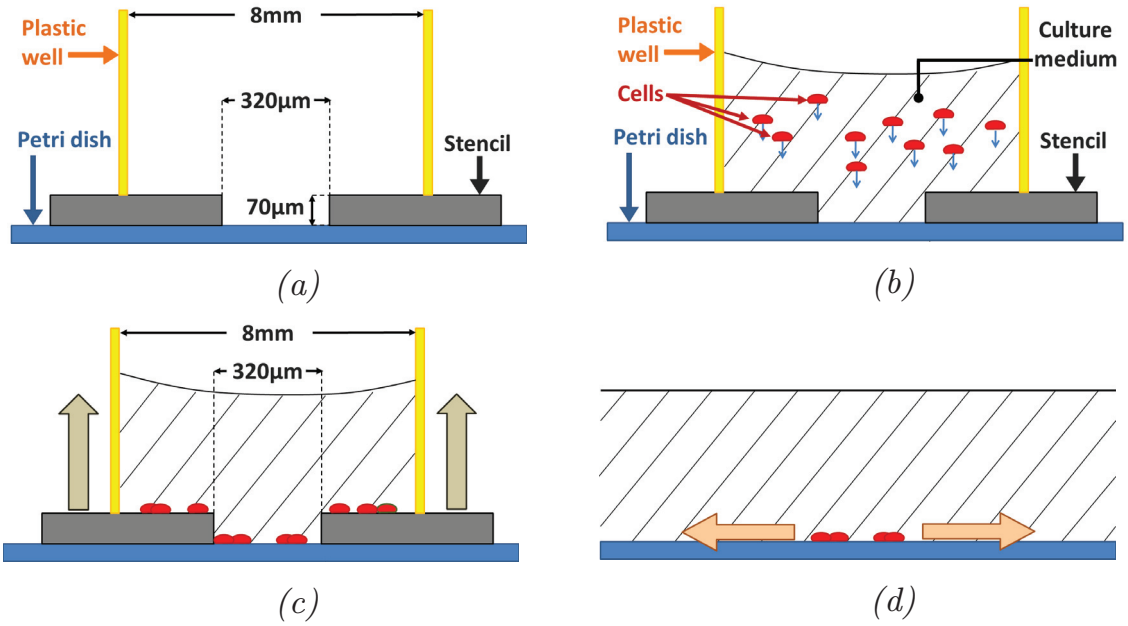


Figure 4.7: Sketch of the patterning procedure.
(a) The PDMS stencil topped with a small plastic well is placed on top of a Petri dish. (b) Cell suspension is deposited into the small well and the cells sediment (c) Once the cells adhere, the stencil and the well are removed. (d) The cells can move freely on the Petri dish. NB: the scales have been modified for clarity.

membrane – since the whole surface of the sample could not be covered – a home-made well made from a cut micropipette-tip was stuck on top of the square so that a droplet of liquid put into that well could contact the sample surface only through the micro-fabricated hole (Fig. 4.7a-b).

We tried glass and plastic surfaces, both untreated or O_2-plasma treated. The best yield was obtained with culture-treated 3.5 cm Petri dishes (BD-Falcon). Indeed, on the other surfaces, the cells were often detached during the stencil peeling-off, probably due to aspiration via medium-entrainment by the chimney walls. Only on the culture-treated dish did the cells adhere strongly enough to resist it, with a designated peeling method.

When we changed the cell suspension density, we observed an unexpected effect: the less concentrated the cells, the longer it took for the suspension to descend the chimney and contact the sample surface. This effect remains mysterious: it might be due to an active modification of the PDMS wall wetting by the cells, or the facilitation, related to the presence of cells, of the evacuation of the air bubble that remains in the hole. An other hypothesis is a fast conditioning of the medium by the cells, which would lead to a modification of its fluid properties. To get

rid of this cell-density dependence and make the experiments in similar conditions for various cell densities, we first dropped 90 µL of culture medium in the home-made well and put the sample under vacuum for 15 min before cell seeding, allowing the prior evacuation of the air from the stencil's hole.

2.5 Cell deposition and stencil peeling off

Cells are seeded in the form of a 60 µL droplet of cell suspension, ranging from 4×10^5 mL^{-1} to 4×10^6 mL^{-1}. The resulting cell number in the 320 µm-wide disk scales linearly with that concentration, from ~ 25 to ~ 250 cells.

Cells are allowed to sediment and adhere on the substrate for 45 min in the culture incubator at 22°C (Fig. 4.7b-c).

To peel off the stencil without detaching the cells from the sample surface, one must first fill the dish with 2 mL of HL5. Then, the joint between the stencil and the well is cut with a surgical blade and the surrounding PDMS is removed smoothly. The well is removed smoothly in turn. Afterwards, the remaining corresponding disk of PDMS can be removed with care and tweezers. During this last step, the numerous cells loosely stuck on the PDMS are in direct contact with the medium and one should take care of not detaching them so that they cannot sediment onto the substrate, outside the patterned colony. Last, the sample is smoothly washed twice with fresh HL5, taking care of preserving the cell colony.

At that point, the sample is ready for observation and should be placed readily under a microscope not to miss the very start of the colony spreading (see Fig. 4.11a).

2.6 Macrofluidic device

For medium perfusion experiments, we need to interface this sample with a fluidic system. The flux will be quite high, allowing for a complete renewal of the sample volume ~ 10 mL in the order of minutes. Balanced in-and-out fluxes in an open system are not very reliable and often lead to microscope drowning. For that reason we first seal the sample with an adapted cover in which an input and an output of calibrated hydrodynamic resistance are plugged, using silicon jointing.

After a 15 min reticulation, the sample is completely filled with fresh HL5 (~ 10 mL) and the input and output tubes respectively connected

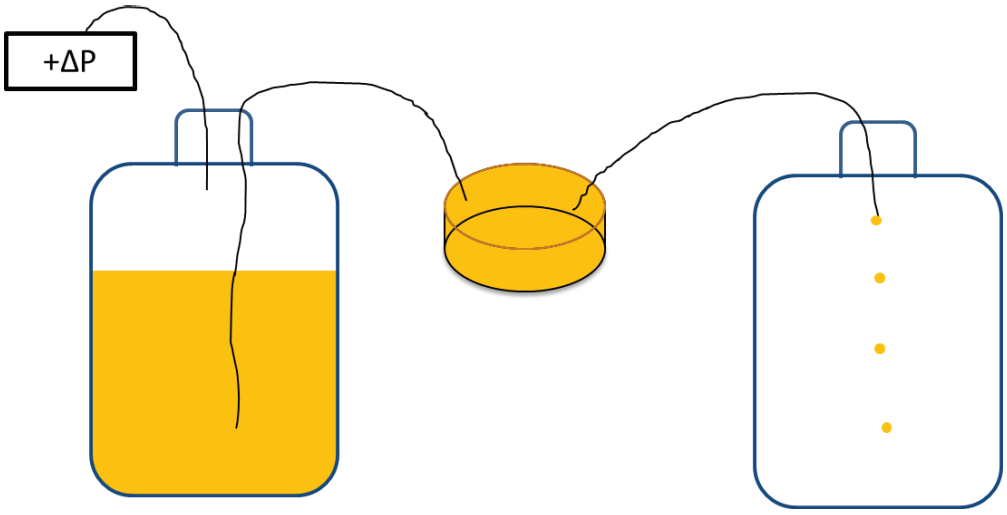


Figure 4.8: Schematic view of the fluidic device. A bottle of HL5 medium is maintained under a controlled overpressure ΔP (left). It is sterilely connected to the sample dish (center), which output is itself connected to a waste bottle right).

to a 1 L-HL5 reservoir and another 1 L empty bottle. All the previous operations must be done under sterile conditions in the hood.

The closed system can be taken under the microscope. The waste is put in the thermalised box while the reservoir is kept outside for convenience reasons. The latter is connected to an automated pressure controller (OB-1, Elveflow, Paris) which is itself connected to a pressure generator (Elveflow, Paris) or a pressured air bottle (Air Liquide, Paris). The pressure controller is automated thanks to adapted software (Elveflow, Paris), allowing in particular to stop the pressure, hence the flow, before the source bottle is empty (Fig. 4.8).

The flow is driven by the pressure difference between the waste bottle – at atmospheric pressure – and the pressure generated in the source bottle. That allows a precise reproducibility of the flow, using always the same tubes and similar sample dishes to keep the same hydrodynamic resistance.

The flux was first calibrated with water pushed through a sample dish. The applied overpressure ΔP and the time t taken to flow 5 mL where measured, leading to calculated fluxes with about 2% errors. From those measurements, the following relation was found:

$$\phi = R_h.(\Delta P + \Delta P_0) \tag{4.12}$$

with $R_h = 0.21 \pm 0.02$ mL.h^{-1}.mbar^{-1} and $\Delta P_0 = -4 \pm 8$ mbar.

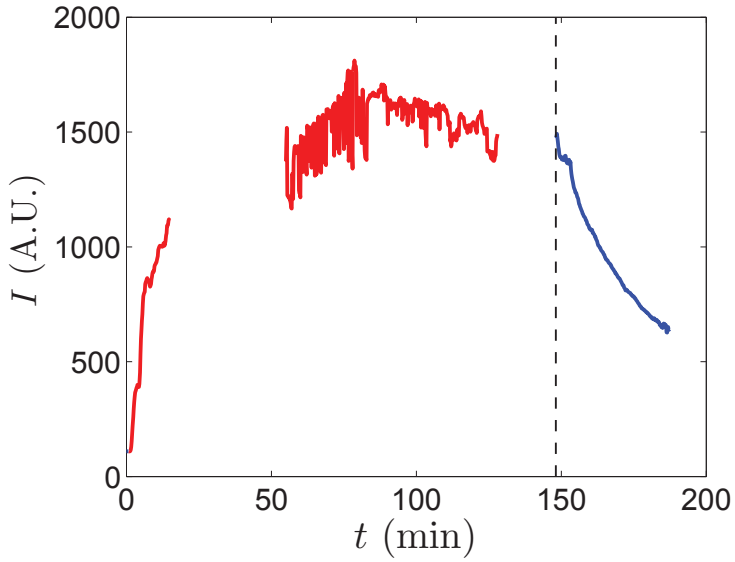


Figure 4.9: Evolution of the FITC fluorescence in time. At $t = 0$, the sample is filled with water. At $t = 50$ s, ΔP is turned on and FITC starts flushing into the sample. At $t = 148$ min, the FITC source is replaced by water, hence rinsing out the FITC from the sample.

Later on, with the experimental samples and HL5 instead of water, these values were found to have slightly changed to $R_{HL5} = 0.33$ mL.h^{-1}.mbar^{-1} and $\Delta P_0 = 11$ mbar.

To check whether such a flux device was able to change efficiently the chemical concentrations close to the bottom surface of the dish, the following experiment was run:

1) A sample was filled with water and connected to the fluidic device.

2) The sample was infused with a FITC solution under $\Delta P = 100$ mbar for 1h30.

3) After a 1 h long pause of the flux, the sample was then infused with water for 45 minutes to rinse out the FITC.

Throughout the experiment, a time-lapse movie of the sample was taken, at 10X magnification, in the FITC fluorescence channel, with a frame every 10 s. The evolution of fluorescence over time clearly shows that such a device allows efficient renewal of the medium close to the surface (Fig. 4.9).

2.7 Imaging and image processing.

We used Low-Magnification Defocused Microscopy as described in **Chapter 2** to image the cells in a 2560x2160 pixels window, at 10X magnifi-

cation – $\delta x = 0.65$ μm – and a frame rate of 3 fps – $\delta t = 20$s.

The cells were located using the Find Maxima based routine, although its accuracy was checked by comparing the results to those obtained from Edge Detection in some experiments, and the cell trajectories retrieved and analysed using home-made Matlab software, as in the rest of this thesis.

In particular, we computed the cell coordinates and velocities in both Cartesian and cylindrical coordinates, allowing to measure many ortho-radially averaged field data.

3 Experimental results

The experimental results from this section are taken from the following available data:

A Series. 4 (respectively 3) successful 2-days long experiments at high density ($N_i \simeq 250$ cells) with AX2 (respectively DH1) cells from March to May 2014.

B Series. 3 experiments with AX2-$aprA^-$ cells from September 2014.

C Series. 3 perfusion experiments from October-November 2014.

D Series. 22 short (> 9 h) experiments at various densities with AX2 cells from December 2014 - January 2015.

E Series. 4 short (> 9 h) experiments at various densities with AX2 cells under fresh HL5 perfusion for 9 h.

F Series. 8 short (> 9 h) experiments at various densities with AX2 cells in highly conditioned medium (HCM).

3.1 Overall view: FKPP-like dynamics

Cell proliferation. In a dilute, homogeneous population of cells, we would expect the number of individuals to grow exponentially, with a characteristic doubling time that is estimated to be $\sim$ 8h for the studied strains AX2 and DH1. According to the Fisher model and to the corresponding simulations by Drasdo *et al.* [102, 103] this exponential growth should hold at short times, when the density is low enough everywhere, before switching to a $N \sim t^2$ growth that can be explained in terms of proliferating rim.

Indeed, we observe such an exponential growth in the experiments, as clearly demonstrated by the semi-log plots of Figure 4.10. By comparing two sets of experiments run with AX2 and DH1 cells (**A series**), we obtain very similar doubling times:

$t_d^{AX2} = (9.4 \pm 0.2\ )$h

$t_d^{DH1} = (9.0 \pm 0.3\ )$h.

Note that these doubling times are very consistent with the usually assumed time of 8 h. Concerning the long-time evolution, it is difficult to state whether the system enters a power-law growth regime or not, and if yes to determine the exponent, even though there is a clear deviation from exponential and the end of both log-log plots suggests it (Fig. 4.10).

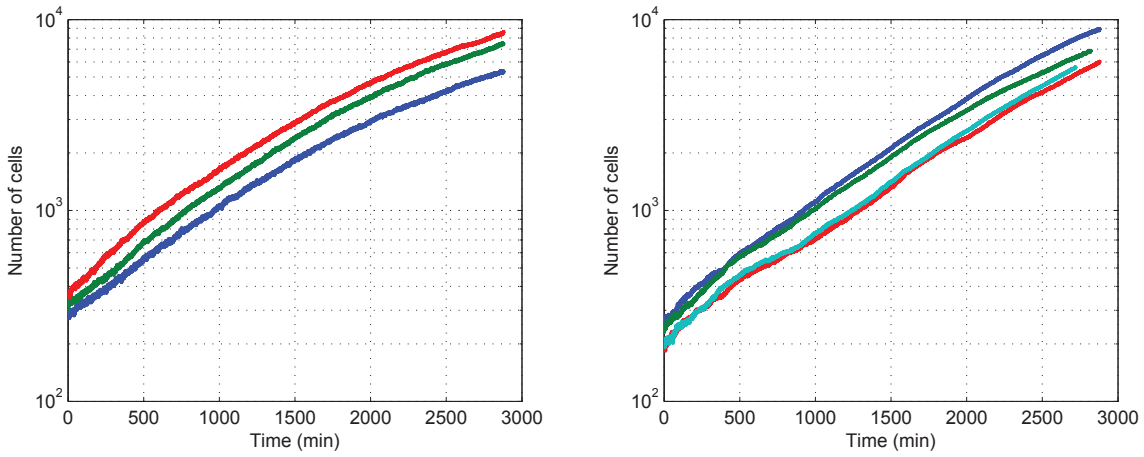


Figure 4.10: Cell growth in DH1 (left) and AX2 (right) colonies. Semi-log representation of the growing number of cells. Each curve is for one experiment from the A series ($n = 3$ and 4 independent experiments respectively).

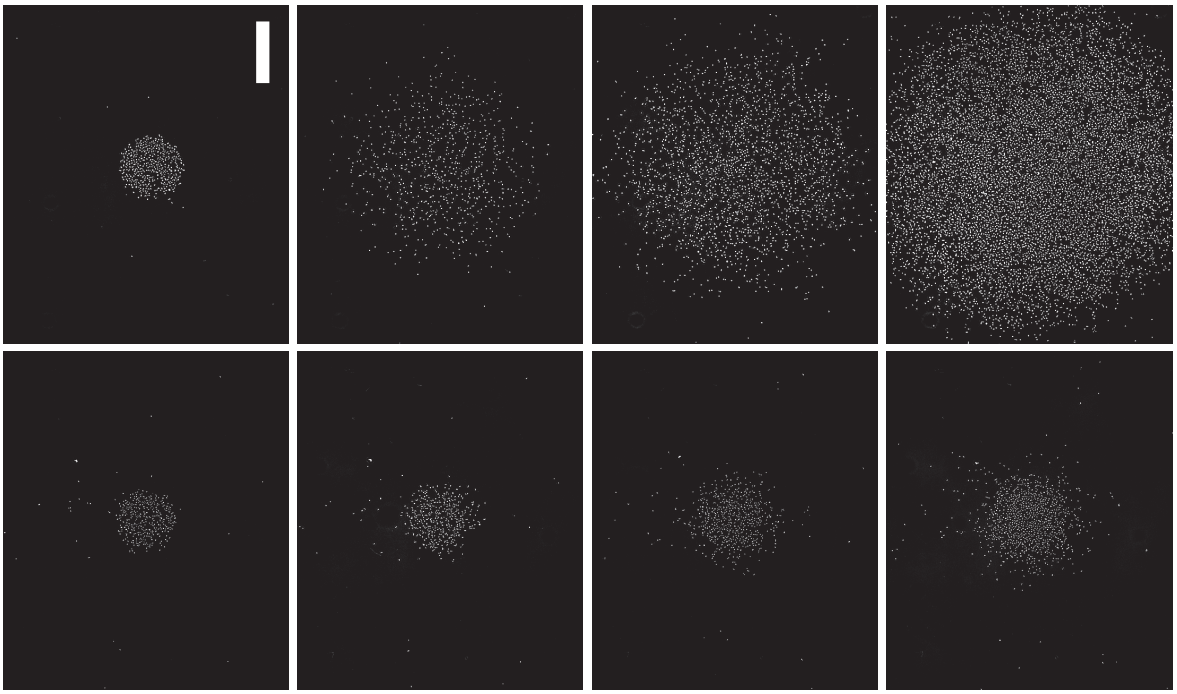


Figure 4.11: Snapshots of the spreading of cell colonies.
Top: AX2 cells at $t = 0$, 100, 300 *and* 600 min *(from left to right).* $N_0 = 327$ *cells initially. Scale bar:* 300 µm. *Bottom: DH1 cells at the same time points,* $N_0 = 324$ *cells. The contrast was enhanced for better visualisation.*

Colony spreading. As specified in the previous section, experiments were run with two different strains whose doubling time was approximately the same. Yet, we had observed in previous experiments that they differed greatly in motility, with about a 4-fold higher D for AX2 than for DH1 – although the D measured from 2014 were about 3 times lower than in 2013, the hierarchy subsisted. This provided us a providential means to decouple motility and growth and study their respective effects.

Our observations fulfil this wish nicely. The AX2 colonies spread faster (Fig. 4.11), as expected, and the shape of the density profiles are very dissimilar from one strain to the other (Fig. 4.12). In the case of AX2, the spreading is fast enough for the density in the center to first decrease before the proliferation becomes dominant, while the density of DH1 colonies increases everywhere from the beginning. In the former, that induces almost linear profiles with very shallow gradients, while in the latter the gradients are steep and the density evolves towards a flat profile with an advancing front. The density reaches a saturation value at $6 - 8 \times 10^5$ cm^{-2} in both cases, but much more quickly in the DH1 colonies.

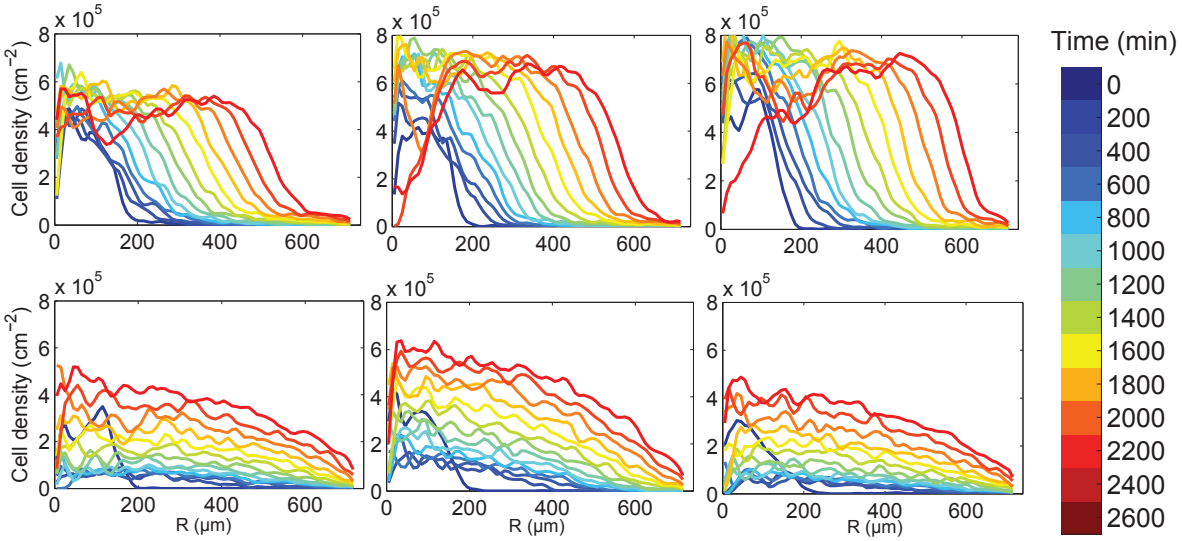


Figure 4.12: Density profiles at various times for 3 DH1 colonies (top) and 3 AX2 colonies (bottom). The hollows at late times on DH1 density profiles are an artifact caused by the high density aggregation that prevents accurate cell detection.
Note: 1 day = 1440 minutes.

Gyration radius. Qualitatively, these evolutions of the density profiles are in accordance to the ideas governing the FKPP equation. To get a better insight in the scaling laws ruling the spreading in both cases, we computed the gyration radius R_g as follows:

$$R_g = \sqrt{\langle r^2 \rangle} \tag{4.13}$$

where r refers to the radial coordinate of a cell with respect to the center of the colony and the brackets $\langle : \rangle$ denotes the average over all the cells in the colony.

The radius of DH1 colonies increases continuously in a way that could match FKPP dynamics, even though the growth exponent looks higher than 1 at long times (see Fig. 4.13a, c).

In the case of AX2 cells (Fig. 4.13b, d), there is a clear transition from a fast to a slower spreading regime. From the log-log plot, it is not possible to estimate scaling for $R_g(t)$ over a broad range of times, but it appears that the FKPP equation will not be enough to account for the colony dynamics. To identify the source of this deviation from the Fisher model, we looked further into the evolution of cell motion over the experiments' duration.

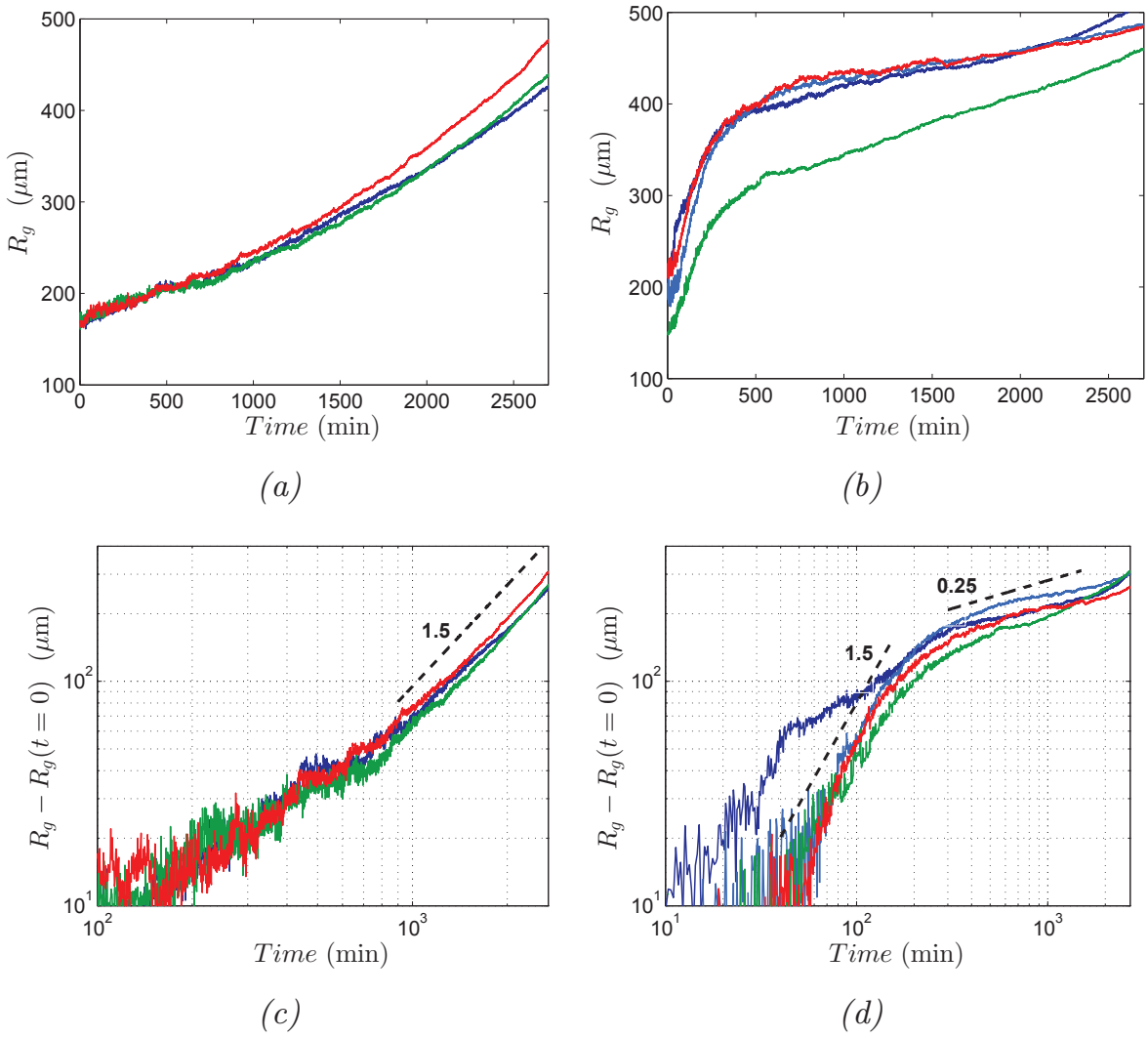


Figure 4.13: Colony radius.
Linear plot of $R_g(t)$ for DH1 (a) and AX2 (b) colonies, and corresponding log-log plots of $R_g(t) - R_g(0)$ (c and d respectively). The dashed lines indicate the approximate slope of these curves. $n = 3$ and 4 independent experiments respectively.

3.2 Long-term variations of the cell motility: a putative QSF effect

Speed and diffusion constant. One of the key points of our data is that we can access single trajectories. Thus, we can not only estimate transport coefficients from a fit with an assumed model, but rather directly measure them in time and space.

First, we studied the diffusion constant, which is supposed to remain constant in the FKPP equation. We define it as $D = \lim_{\delta t \to +\infty}(\frac{\text{MSD}}{4.\delta t})$, where the limit is computed at $\delta t = 22.3$ min since we cannot access infinite times. This is an average over the entire colony, made in successive 67-minutes-long time windows, which defines the time resolution of this

measurement.

To get a better resolved quantity, we also measured the instantaneous "speed", defined as $v = \frac{\delta x}{\delta t}$, where δx is the displacement of a cell between two successive frames, and $\delta t = 20$ s the time between two frames. Although v is somehow ill-defined in this way, as we saw in Chapter 2, it still gives a valuable estimate of the short-time speed of the cells.

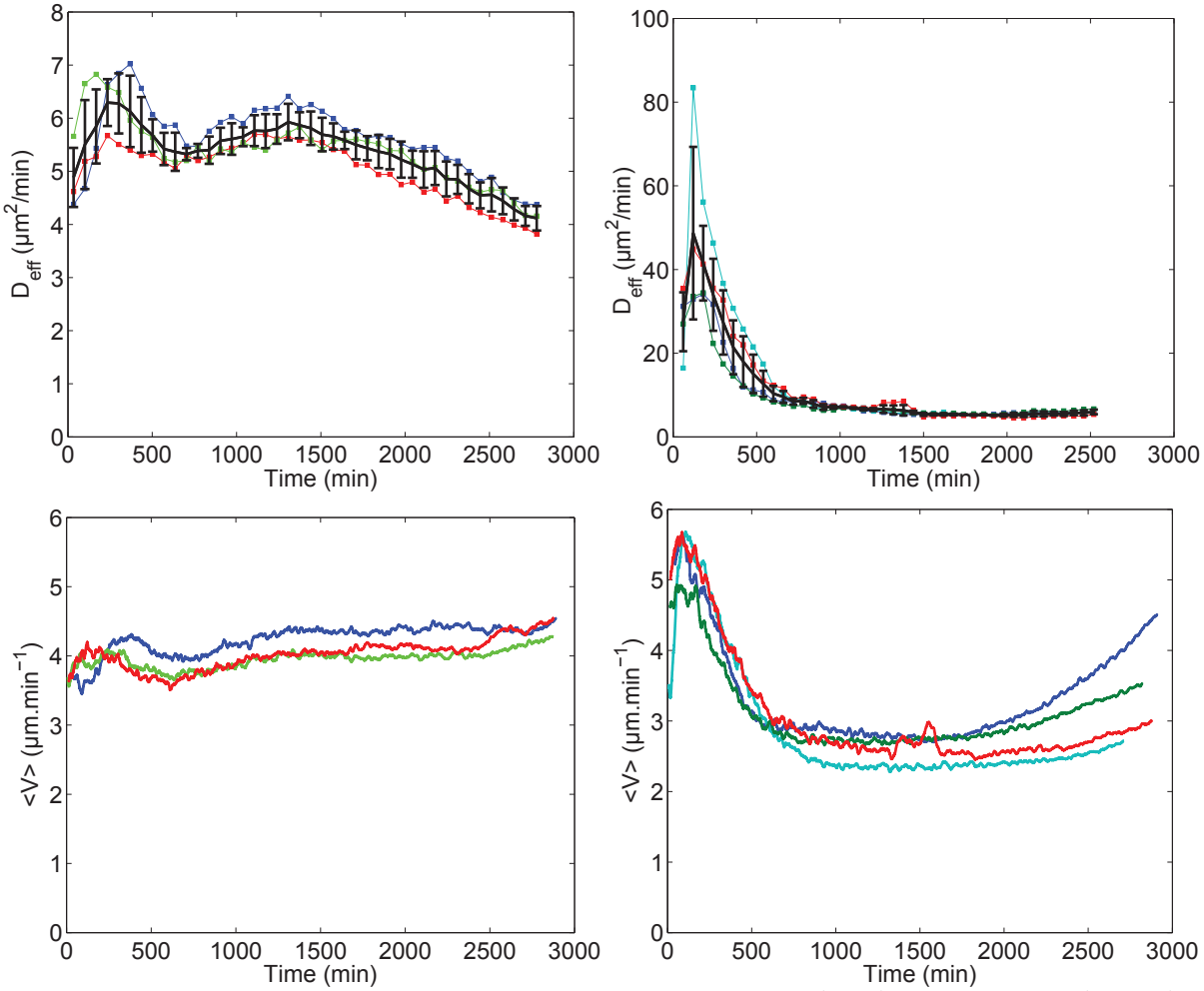


Figure 4.14: Colony averaged motility for DH1 (left) and AX2 (right) cells.
Top: diffusion constant. Each color is for one experiment, the black curve shows the average over all experiments and standard deviation.
Bottom: instantaneous colony-averaged speed. Each curve is for one experiment.

Looking at their colony averages throughout the experimental time, one feature appears clearly: they are not constant at all, especially in AX2 colonies (Fig. 4.14). In the latter, both $\langle v \rangle$ and D increase until a peak at $t \sim 90$ min, then decrease for about 10 h to reach a low plateau value.

Compared to AX2, the absolute values do not vary much in the case of DH1. Yet the shape of AX2's initial peak is recalled in both D and

$\langle v \rangle$ curves, and D exhibits a slow but continuous drop at $t > 1300$ min. The inter-strain difference in motility is clear on the compared values of D while both speeds are similar. Therefore the difference should occur from a difference in persistence, with more random trajectories in the case of DH1. We can roughly estimate an "average persistence time" τ_p, assuming a simple correlated random walk, through the use of $\tau_p = \frac{2D}{v^2}$. We find $\tau_p \simeq 0.6$ min for DH1 and $\tau_p = 1.2 - 3$ min for AX2 in time.

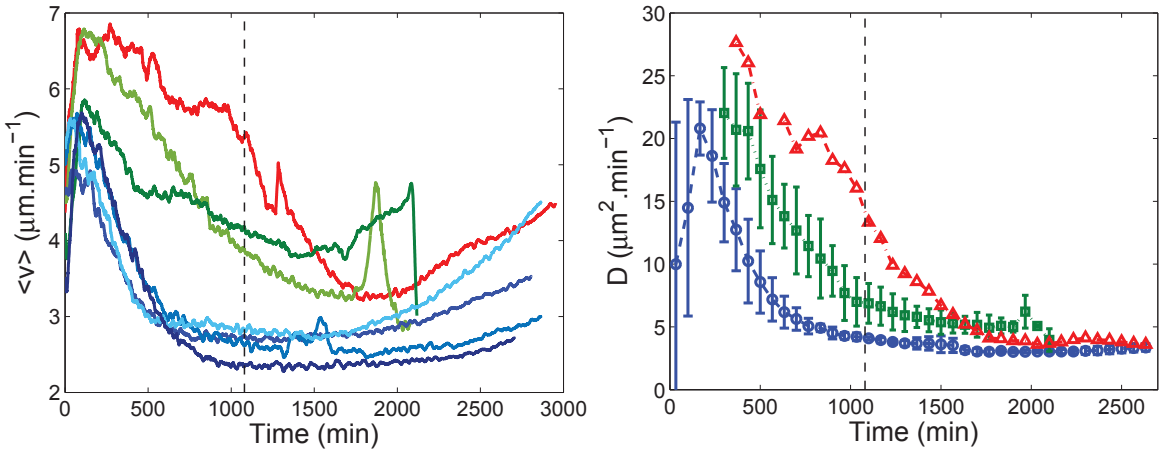
Perfusion experiments. The observed long-term decrease of the motility is not a surprise for the careful reader, as we saw in Chapter 3 that *Dictyostelium* cells self-regulate their motility through a secreted QSF. To experimentally check whether such an effect could be at stake here, a fluidic device was used to continuously infuse the samples with fresh medium. Indeed, it has been shown [18] that medium renewal could efficiently suppress its effect.

At a $\Phi = 20$ mL.h^{-1} flux, corresponding to a renewal rate of 200%.h^{-1} since the sample volume is about 10 mL, the motility of the cells still decreases over time, although this may be a little slower (Fig. 4.15). At $\Phi = 100$ mL.h^{-1}, this decrease was significantly reduced over 18 h. Then the flux was stopped, leading to a decrease of D towards its low plateau.

The same kind of experiments was redone (E series) on shorter time scales. In these experiments, the motility is maintained high throughout the duration of the perfusion (9 h) and then it dropped fast (see Fig. 4.19c in **Section 3.3** for a plot of the cell speed).

These results, together with HCM experiments (see the Section 3.3), show that a QSF – presumably the same as described in Chapter 3 – is responsible for the long-term decrease of both the speed and the diffusion constant of the cells.

Thus, the cell motility is time-dependent, which constitutes a first deviation from the simple Fisher model. There is experimental evidence that the motility is down-regulated through a secreted QSF, and this effect is noticeable on long time-scales. Yet it does not explain the peak of activity during the early spreading of the colonies. This is why we investigate further these short-time dynamics in the following section, with a focus on AX2 cells that are the most exemplary of this behaviour.



(a) *(b)*

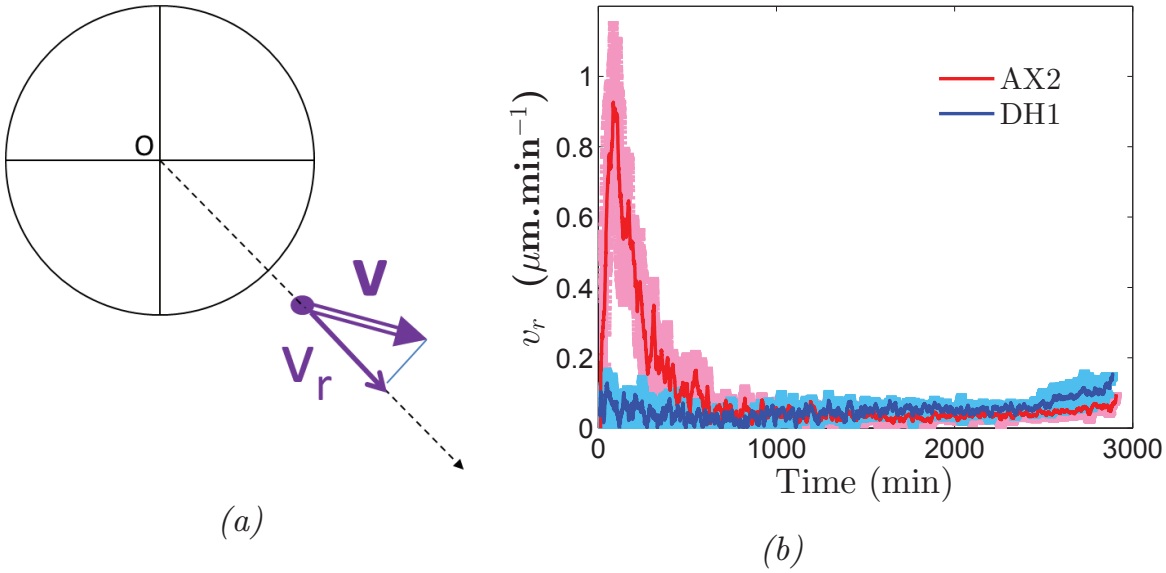
Figure 4.15: Perfusion of FM slows down the decrease of motility.
(a) Average cell speed without perfusion (4 experiments, blue), with perfusion at $20\,\mathrm{mL.h^{-1}}$ *(2 experiments, green) and* $100\,\mathrm{mL.h^{-1}}$ *(1 experiment, red). (b) Diffusion constant (taken at* $\delta t = 15\,\mathrm{min}$*) for the same experiments (circles, squares and triangles respectively). Mean and standard deviation for each condition.*
The dashed vertical lines mark $t = 1080\,\mathrm{min}$*, when the perfusion stops in the high flux case.*

3.3 Density-dependent polarised motion in the short-time regime.

Outwards motion. To that point, we have implicitly assumed that the cell motion is isotropic. In terms of Cartesian coordinates, this is true: there is no symmetry breaking, either at the scale of the colony or at the single cell level.

This last point can be studied by computing the average *velocity vector*. As such, it constitutes a polar order parameter that bears information about the presence of directed bias in the motion. If it is isotropic, $\langle \mathbf{v} \rangle = \mathbf{0}$, while $\|\langle \mathbf{v} \rangle\| \rangle 0$ means that, in average, the cells move in the direction pointed at by $\langle \mathbf{v} \rangle$. The amount to which they are polarised can be quantified by the "directionality index" $I = \frac{\|\langle \mathbf{v} \rangle\|}{\langle \|\mathbf{v}\| \rangle}$, even though we will see later more rigorous quantifiers based on the full angular distribution.

Now, when the average is taken in *polar* coordinates, a clear anisotropy emerges: the radial coordinate of the velocity exhibits an important peak, especially pronounced in the case of AX2, whose timing correlates with the higher initial spreading rate (Fig. 4.16).



(a) *(b)*

Figure 4.16: Radial velocity.
(a) The radial velocity v_r is defined in cylindrical coordinates with respect to the center of the colony O. (b) Average radial component of the velocity for AX2 (red, peaked) and DH1 (blue, flat) colonies. The averages have been first computed in each colony, then over all experiments. The light background zones represent the standard deviation over the experiments ($n = 4$ experiments for AX2 and 3 for DH1).

As the radial symmetry is conserved, all the quantities are only functions of r and t, and therefore can be displayed as spatio-temporal colour maps. In Figure 4.17, the maps of $v_r(r,t)$ and $I_r(r,t) = \frac{v_r}{\|\mathbf{v}\|}$ show that during the 200 first minutes, the cells on the edge of the colony move in average outwards. This effect appears also in DH1 colonies but both its amplitude and duration are smaller, explaining why their effect are less noticeable at the scale of the colony.

The effect of cell density. As there is no external cue, this effect occurs from the colony itself, hence from the interactions between the cells. Therefore there should be an effect of the cell density. That is why a series of experiment (**D series**) with various initial cell numbers was run. From these experiments, it is clear that the spreading rate increases with the cell number – or cell density, since the initial diameter is not changed (Fig. 4.18). The difference in spreading occurs at very short times: at $t = 100$ min the profiles are already well separated, and the separation does not increase a lot until $t = 150$ min (Fig. 4.18a-b).

The plot of the gyration radius R_g (Fig. 4.18c) confirms this tendency, and shows that the peak in spreading rate is very correlated to that of

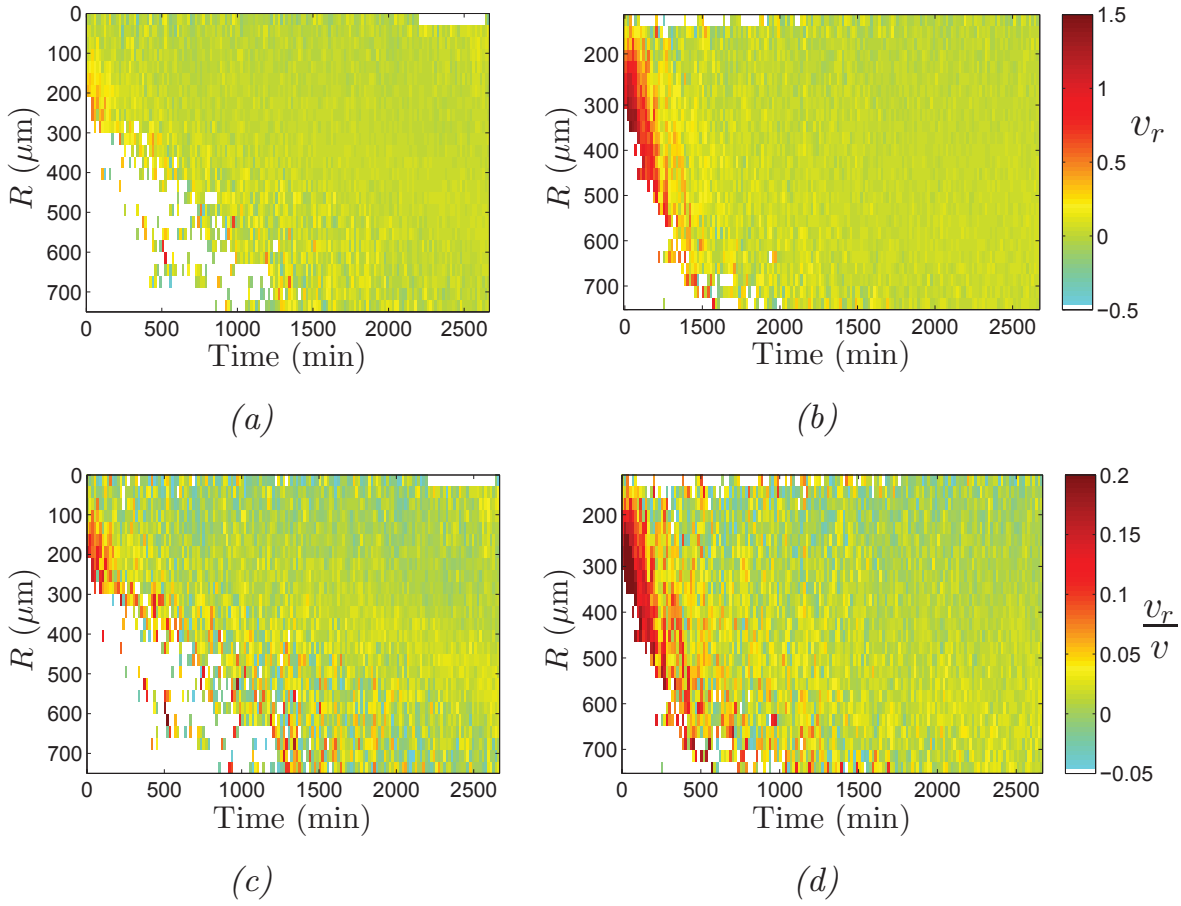


Figure 4.17: Directed cell motion on the edge.
(a-c) DH1 colonies, (b-d) AX2 colonies. (a-b) Spatio-temporal representation of the local average radial velocity (left, in $\mu m.min^{-1}$*) and (c-d) directionality index* $\frac{v_r}{v}$*. The white areas are zones with no detected cells. During the 200 first minutes, cells on the edge of the spot move in average towards the outside of the colony.*

the radial velocity. Interestingly, the increase of $\langle v_r \rangle$ seems to be borne by the same curve for all initial cell numbers N_0, and only the time it stops increasing – *ie* the peak's timing – seems to depend on N_0 (Fig. 4.18d). Thus one can define $t^*(N_0)$ as this peak position, which scales almost linearly with N_0 (see **Section 5.2** for more details). This last finding must be taken with care, due to the error on $\langle v_r \rangle$: not shown for the sake of clarity, they are of the order of 20% and stem mainly from the inter-experimental variability, and they could affect the determination of t^*. Yet, it seems to be a particular feature of this system and we will come back to it later to try to understand its origin.

Thus, the fast early spreading is well explained by an outward-directed motion, which originates from cell-cell interactions yet to be identified. Two natural candidates stand out from the current knowledge: long-

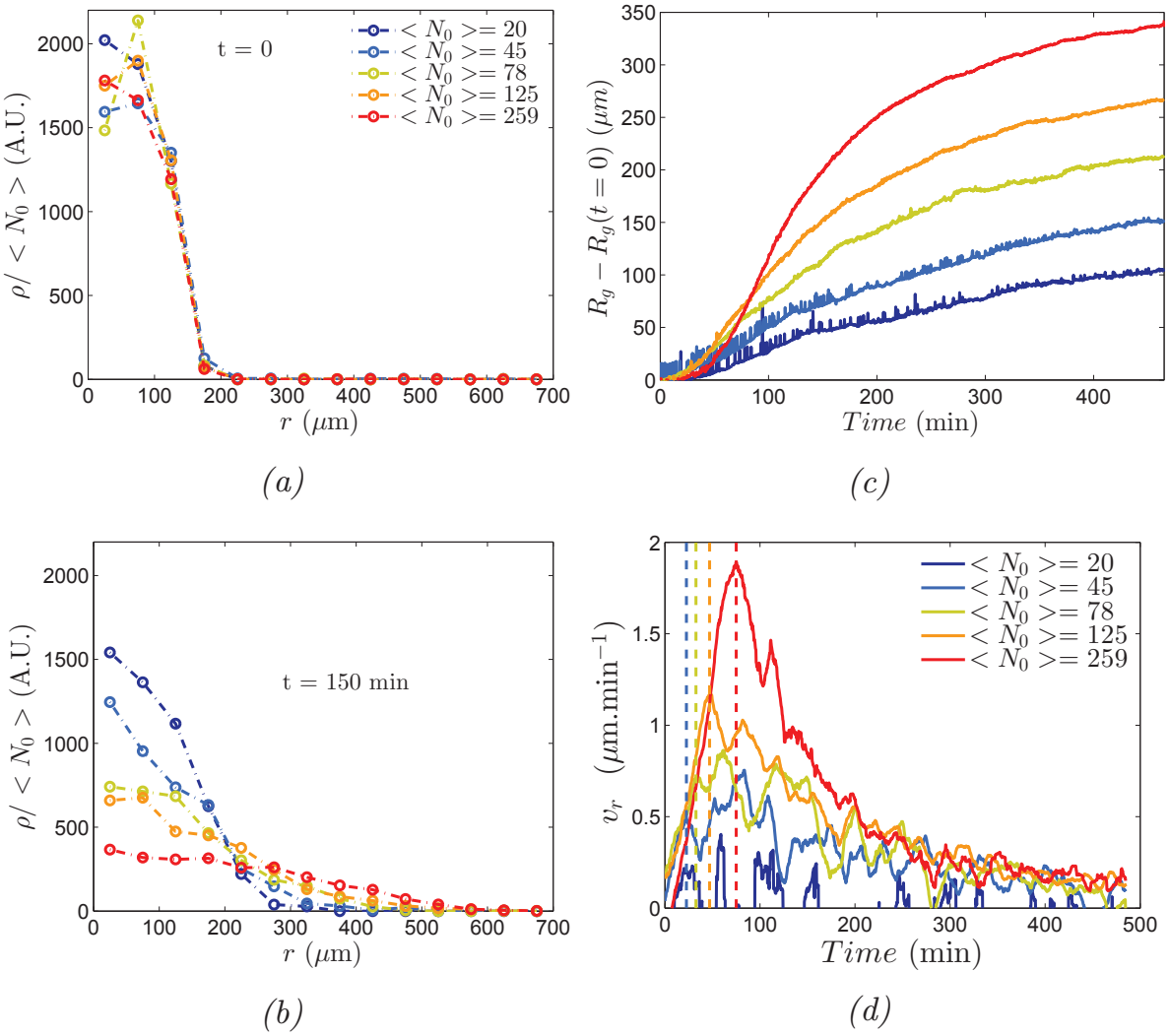


Figure 4.18: Density-dependence of the spreading.
(a) Normalised density profiles at $t = 0$ for various starting cell numbers N_0. (b) Same at $t = 150$ min. (c) Radius increase. (d) Radial velocity. The vertical dashed lines denote the position of $t^*(N_0)$, the peak timing for each cell density.

range repulsion through chemotaxis, or a short-range effect due to collisions. We first focused on the former.

Chemo-repulsion by AprA is not required. It has been shown recently [22] that AprA is an endogenous chemorepellent in vegetative *Dictyostelium discoideum*. It seemed natural to check whether cells that do not secrete this protein will still exhibit the fast spreading dynamics and outward escape in a colony experiment. A series of experiment (**B series**) was run with AX2-*aprA*$^-$ cells. Although the motility of individual cells is different from that of our wild-type AX2 strain, as already mentioned in Chapter 3 of this thesis, the features of directed

escape are still observed in these experiments (Fig. 4.19a).

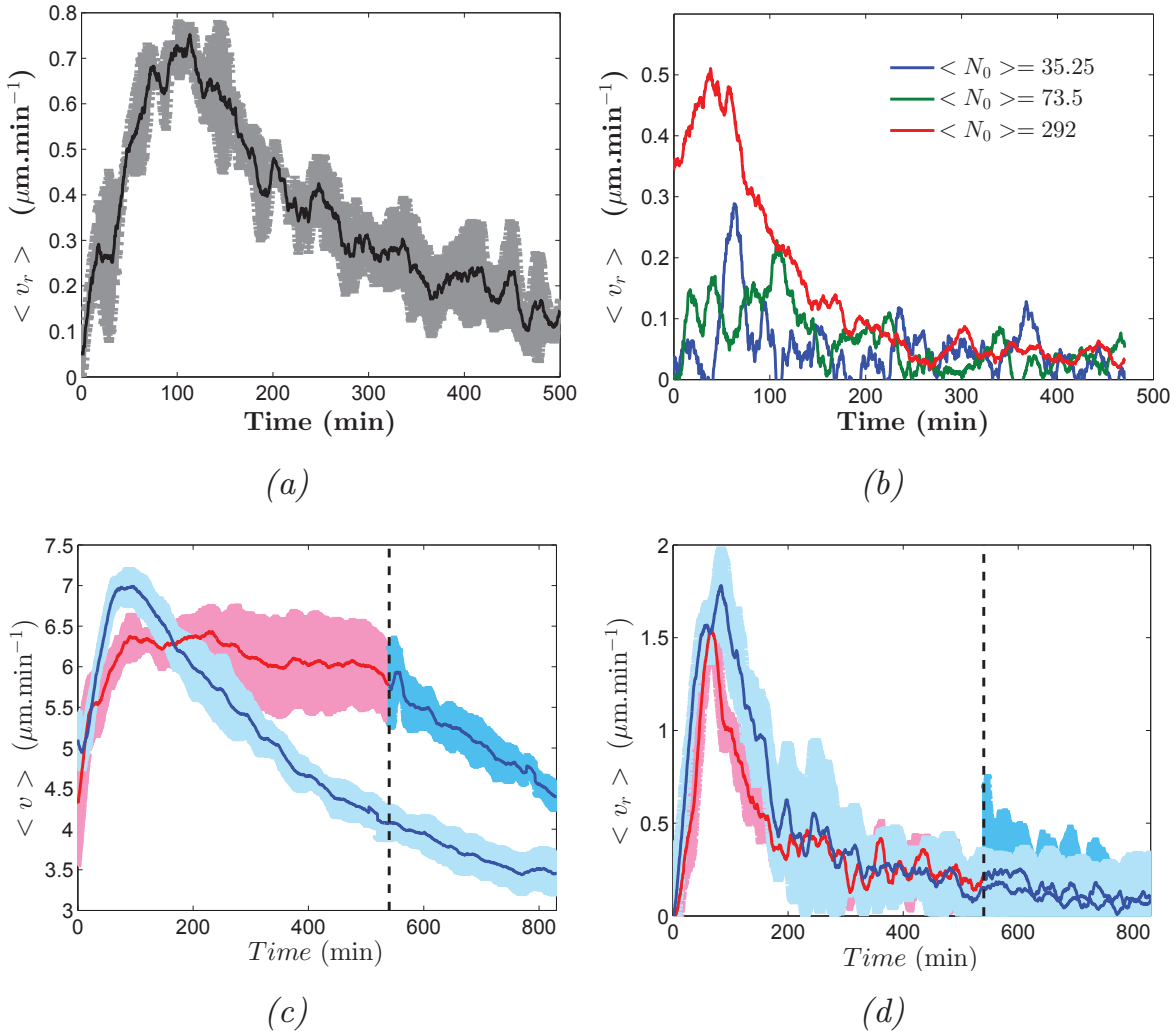


Figure 4.19: The collective effect does not rely on large-scale chemical interactions.
(a) Radial velocity for AX2-aprA$^-$ cells at high initial density ($\langle N_0 \rangle$ = 312). The cell speed is $v \simeq 5\,\mu\text{m.min}^{-1}$. (b) Radial velocity for colonies in HCM at various N_0. The cell speed is $v \simeq 3\,\mu\text{m.min}^{-1}$. n =4, 3 and 3 experiments respectively. (c) Colony-averaged cell speed as a function of time for AX2 colonies perfused with fresh medium for 9 h (red) or without perfusion (blue). The speed remains high until the flow is stopped (dashed vertical line), then it starts decreasing. Each curve is an average over four distinct experiments, the light background represents standard deviation ($\langle N_0 \rangle \simeq 250$ in both cases).
(d) Colony averaged radial velocity for both perfused (red) and non-perfused (blue) colonies, from the same experiments as in (c). v_r remains unaffected by the perfusion.

Thus, AprA does not seem to play a major role in this effect. This

result does not rule out the possibility of chemorepulsion, certainly, but it undermines it to the extent that if chemorepulsion was at stake, the presence or absence of AprA should be noticeable [22]

Perfusion experiments. Let us recall that in the perfusion experiments of the **E series**, the average cell speed and diffusion constant are maintained constant along the perfusion duration, showing that the secreted molecules are efficiently rinsed out. Any chemorepellant activity should be suppressed in the same way, but the peaks in radial velocity and spreading rate hold (Fig. 4.19c, d). It demonstrates that the outward escape does not rely on a large-scale chemical effect.

The outward motion survives high conditioning of the medium. Last, experiments in highly-conditioned medium (HCM) show the same tendency. In a simple view on such a medium, every secreted factor is expected to be present in a huge concentration. Supposing that the cell chemosensing systems saturate above some threshold value, any putative chemorepellent should be above this saturation. It would lead to a screening of the chemorepellent field, thus suppressing the outward escape in case it is due to chemotaxis. Although the motion is again different from that of the wild-type experiments in fresh medium, witnessing the presence of QSF and so the high level conditioning of the medium, the outward motion again survives (Fig. 4.19b).

These three different experiments together demonstrate that the outward escape of the cells is not due to a long-range chemical effect. Therefore, it has to originate from local interactions between neighbouring cells, likely contact interactions.

In the following, before exploring various modelling approaches that will give insight into the nature of these contact interactions, I would like to first describe in more details the spatio-temporal dynamics of the colonies, and especially the appearance of polarity.

3.4 Polarisation dynamics

A true polarisation signal. To better understand the rise of directionality in the dense colonies, we examined the spatio-temporal dynamics of various observables. The maps of density and radial velocity are reminiscent of already described phenomena. The speed $v(r,t)$ itself increases in the zones of high v_r (Fig. 4.20). Yet, without anisotropy, it would not yield such high values of v_r. To check the origin of anisotropy,

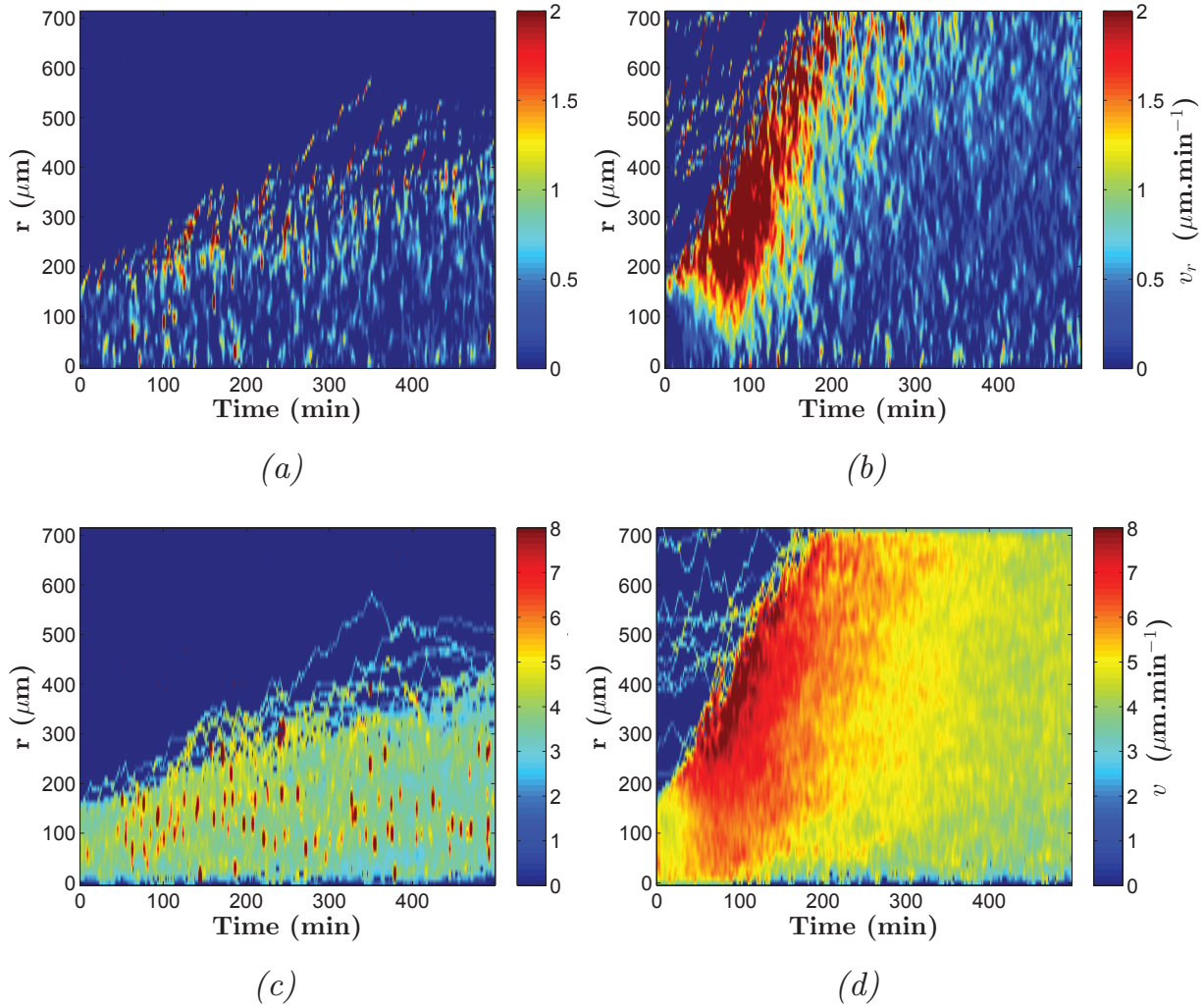


Figure 4.20: Spatio-temporal dynamics of the polarisation
Only $\langle N_0 \rangle = 20$ [left column] and $\langle N_0 \rangle = 259$ [right] are displayed here. (a, b) Radial velocity (c, d) Speed (norm)

we looked into the angular density function $\mathbb{P}(r, \varphi, t)$ where φ denotes the angle of the velocity vector with respect to the radial direction e_r. It can be decomposed in angular harmonics (Fig. 4.21):

$$\mathbb{P}(r, \varphi, t) = \rho(r, t) + \mathbf{p}(r, t).\mathbf{u} + \mathbf{q}(r, t) : (\mathbf{uu} - \mathbf{I}/2) + \psi(r, \varphi, t), \quad (4.14)$$

where $\mathbf{u} = (cos(\varphi), sin(\varphi))$ and ψ contains higher order harmonics. ρ is simply the cell density. The polarity $\mathbf{p}$ characterises the strength of a mono-oscillation perturbation: its adimensioned version $\mathbf{P} = \mathbf{p}/\rho$ is a polar order parameter, while $\mathbf{Q} = \mathbf{q}/\rho$ characterises the nematic order through a symmetric two-oscillations perturbation.

In concrete terms, we compute them as $\mathbf{P} = 2\langle \mathbf{u} \rangle_\varphi$ (the factor comes from the summation of harmonics -1 and +1) and $(Q_r, Q_\varphi) = 2\langle (\cos(2\varphi), \sin(2\varphi)) \rangle_\varphi$

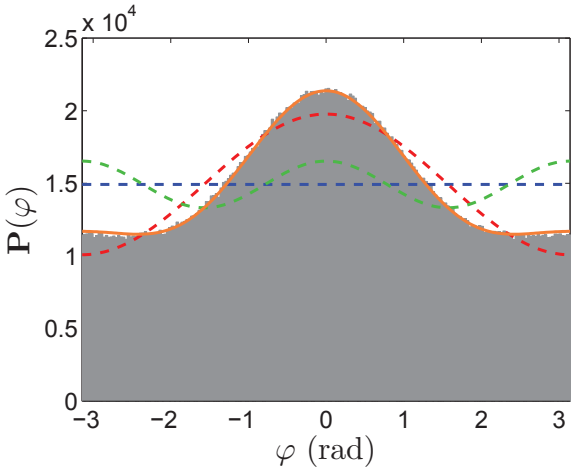


Figure 4.21: Example of angular density function $\mathbb{P}(\varphi)$ and its decomposition in angular harmonics.
Here $\mathbb{P}(\varphi) = P_0(1 + p_1\cos(\varphi) + p_2\cos(2\varphi))$ is polar: $P_0 = 1.5 \times 10^4 AU$, $p_1 = 2\langle\cos(\varphi)\rangle = \frac{1}{3}$ and $p_2 = 2\langle\cos(2\varphi)\rangle = \frac{1}{10}$. The dashed lines show the 0^{th} (blue), $0^{th}+1^{st}$ (red) and $0^{th}+2^{nd}$ (green) harmonics, the orange solid line is the sum of the three.

to simplify. We also measured the 'speed anisotropy index' $\langle\|\mathbf{v}\|\mathbf{u}\rangle\varphi$, which tells if the speed itself is anisotropically distributed. Here, the notation $\langle:\rangle_\varphi$ refers to the average over all possible orientations φ. We find that only P_r takes significant values (Fig. 4.22a-b), indicating the order arising is strictly polar – even though it might be shaped more finely by higher harmonics – and controlled by the angular distribution of the cells, not by an anisotropic speed.

As a check, we computed $v_r(r,t) = \|\mathbf{v}(r,t)\|.P_r(r,t).\rho(r,t)$, which shows a very good agreement with the measured field (compare Fig. 4.22c-d to 4.20a-b).

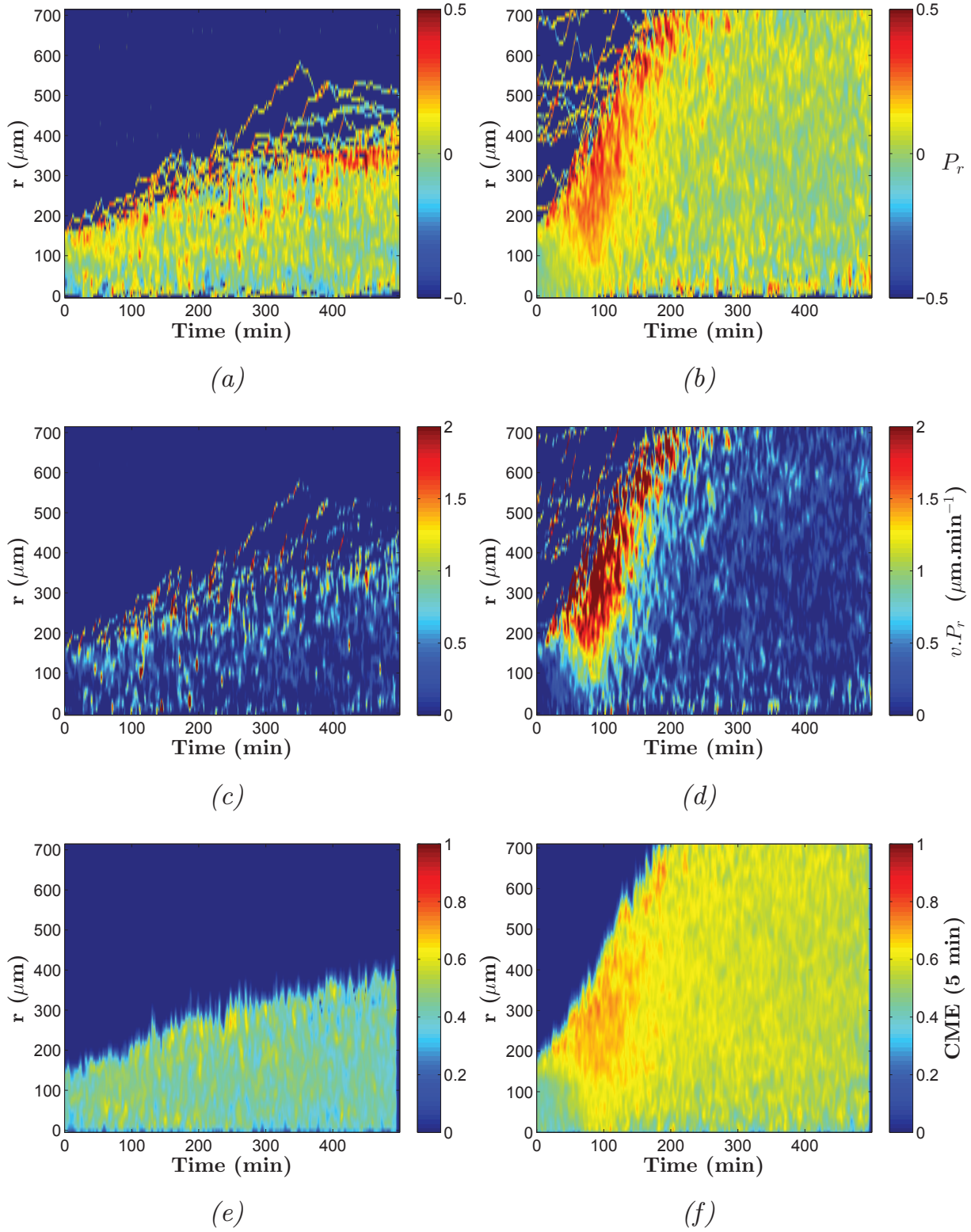


Figure 4.22: Spatio-temporal dynamics of the polarisation (2). (a, b) Polarity (c, d) Computed radial velocity $v_r^{comp} = v.P_r$ (e, f) CME measured at 5 min.

Inward propagating directionality signal. A more careful look into the v_r dynamics for $t < t^*$, *ie* before the peak in v_r shows that the onset of directionality happens at the very edge of the colony, at $t \simeq 0$ (Fig. 4.23). Then, the signal propagates inwards at constant speed, while it builds up on the edge. Despite the noisy nature of these maps, which prevents us to make really quantitative assertions on them, notice that:

(i) the signal v_r expands outwards due to colony spreading, but more importantly it propagates towards the center of the colony with a propagation speed of $c \sim 1$ µm.min^{-1}, which is of the same order as the cell speed;

(ii) again, the maps for various N_0 all look similar to each other, with the only difference being t^*, which corresponds to the end of the signal propagation and amplification.

Increase in persistence. Last, we studied the persistence of the trajectories. To that end, the velocity autocorrelation function is the canonical intermediate to access parameters such as the persistence time in a simple correlated random walk, or the proportions of the modes in a multi-modal motion (see **Chapter 2**). Its drawback is the same as the diffusion constant's: it requires an average over several long enough trajectories to be usable, thus reducing the accuracy of any measurement. To get resolved spatio-temporal dynamics of the local persistence, we computed the "coefficient of movement efficiency" (Fig. 4.24), or CME, which we define, for a given time interval Δt, as the ratio between the end-point distance $\|\mathbf{r}(t+\Delta t) - \mathbf{r}(t)\|$ and the total distance travelled during Δt (see also **Chapter 2**):

$$CME_{\Delta t}(t) = \frac{\|\mathbf{r}(t+\frac{\Delta t}{2}) - \mathbf{r}(t-\frac{\Delta t}{2})\|}{\int_{t'=t-\frac{\Delta t}{2}}^{t+\frac{\Delta t}{2}} \|\mathbf{v}(t')\| dt'}. \tag{4.15}$$

Measuring it over temporal sliding windows, we could build its spatio-temporal map for all N_0 and various Δt (Fig. 4.22e–f, only shown for $\Delta t = 5$ min). They exhibit clearly higher values maintained over longer Δt in the zones of high polarisation. This suggests that the rise of polarisation is linked to an increase of the single cells' persistence. We measured the velocity autocorrelation functions in 67 min-long time windows. As in other experiments it is best fitted with a bi-exponential

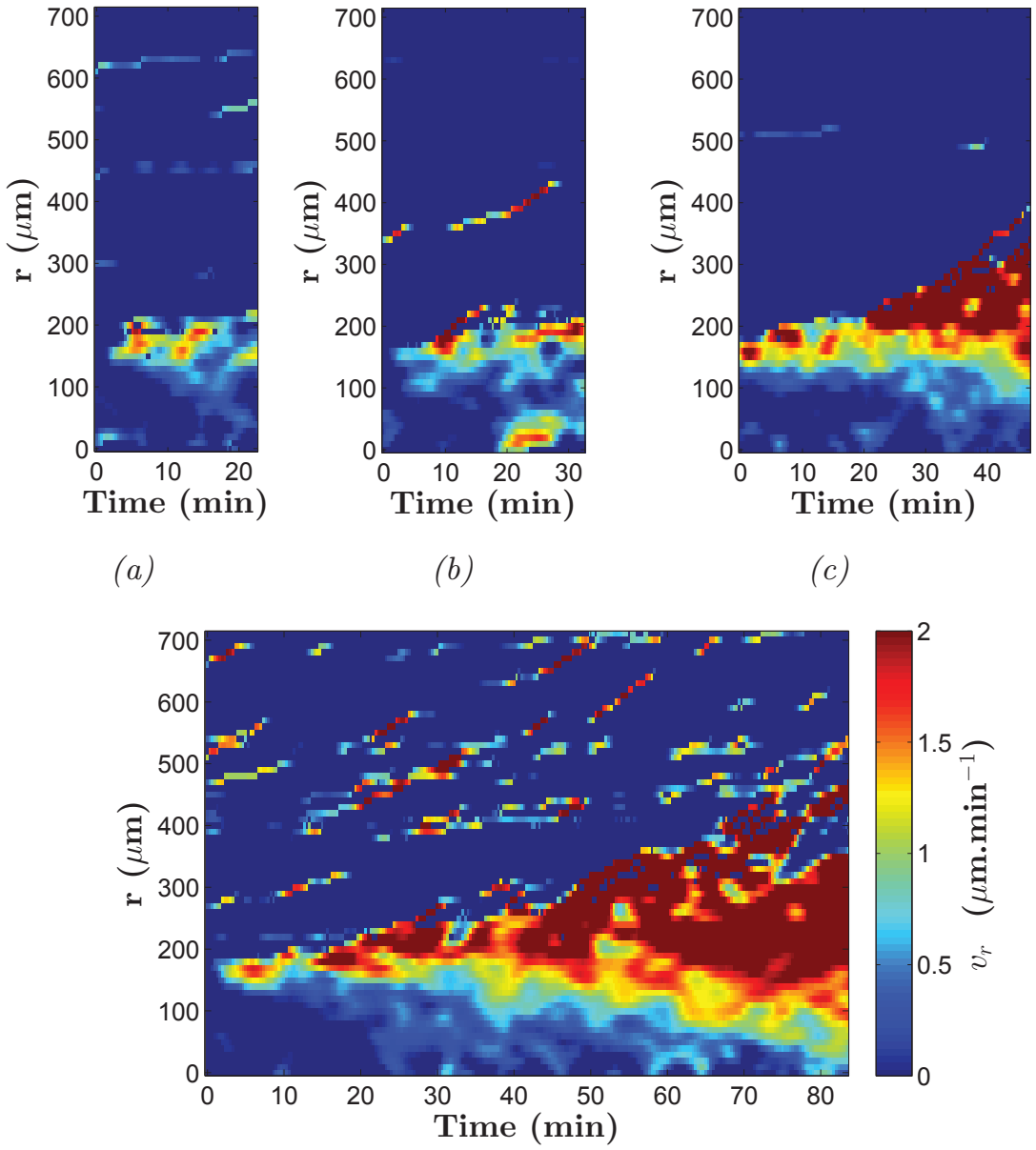


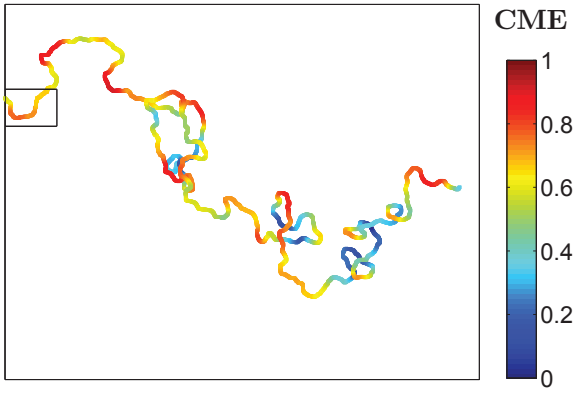
Figure 4.23: Polarity wave.
Zoom in v_r maps for (a) $\langle N_0 \rangle = 45$, (b) $\langle N_0 \rangle = 78$, (c) $\langle N_0 \rangle = 125$ and (d) $\langle N_0 \rangle = 259$, with $t \leq t^*(N_0)$.

decrease rather than a single exponential (Fig. 4.25, see **Chapter 2, Sections 3.3** and **4.2**):

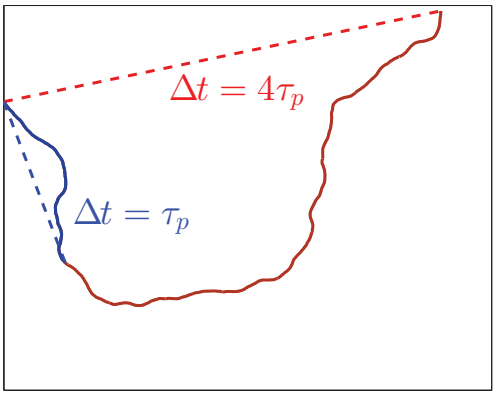
$$C(t) = \phi_1 \exp(-\frac{t}{\tau_1}) + \phi_2 \exp(-\frac{t}{\tau_2}) \tag{4.16}$$

When $\tau_1 \in [1;3]$ min and $\tau_2 \geq 7.5$ min, the fit's accuracy is really good (Fig. 4.25a). Again, we fix two of the parameters – $\tau_1 = 2$ min and $\tau_2 = 10$ min – which allows us to sample $\phi_1(t)$ and $\phi_2(t)$ in a consistent

(a)



(b)



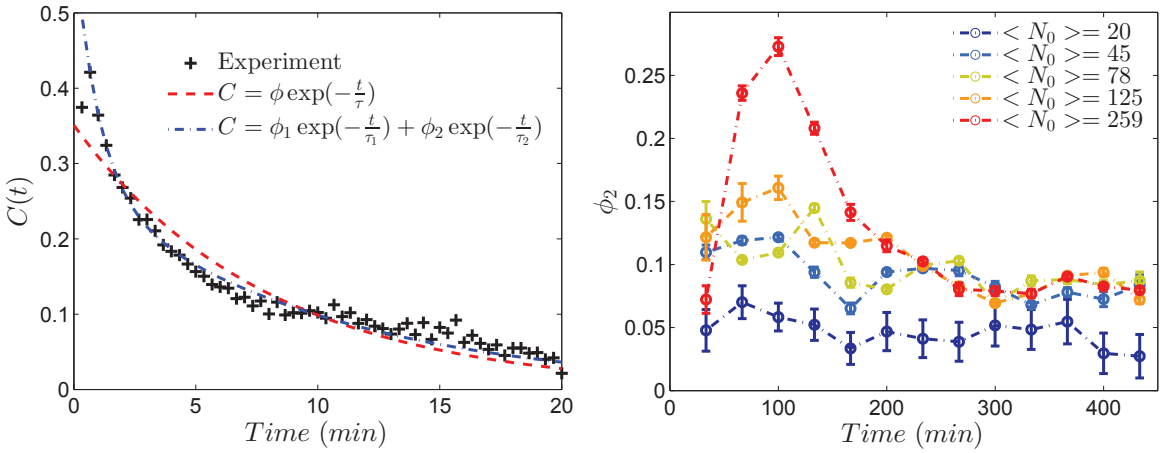
(c)

Figure 4.24: Coefficient of Movement Efficiency.
(a) Trajectory of an active Brownian particle with colour-coded instantaneous CME computed at $\Delta t = \tau_p$. (b) Same trajectory with CME computed at $\Delta t = 4\tau_p$. (c) Zoom in the trajectory (rectangle in (a) and (b)), for $t = 0 - 4\tau_p$. The CME is computed as the ratio of end-point distance over Δt (dashed lines) on the total path length (solid lines).

way over successive time windows. $\phi_2(t)$ increases with N_0 and has a maximum at about t^*, which confirms that the persistence is enhanced when the polarisation occurs, seemingly because then the cells spend more time in a very persistent mode (Fig. 4.25b). As this effect relies on local cell-cell interactions and it improves the large scale motion of the cells, we term it **"Contact Enhancement of Locomotion"**.

3.5 Summary

Let us sum up our main experimental results before we tackle the models we used to handle them. We observed the spreading of AX2, DH1 and AX2-*aprA*$^-$ colonies in non-renewed FM, and AX2 colonies with FM perfusion or in HCM. Despite the differences inherent to these various situations, we could notice some features that are common to them all.



(a) *(b)*

Figure 4.25: Contact Enhancement of Locomotion.
(a) Autocorrelation of the velocity direction. The red dashed line is a mono-exponential fit $C(t) = \phi \exp(-\frac{t}{\tau})$, *which yields* $\tau \simeq 5$ min, *while the blue dash-dotted line is a bi-exponential fit* $C(t) = \phi_1 \exp(-\frac{t}{\tau_1}) + \phi_2 \exp(-\frac{t}{\tau_2})$ *with fixed* $\tau_1 = 2$ min *and* $\tau_2 = 10$ min $(\langle N_0 \rangle = 259$ *cells).*
(b) $\phi_2(t)$ *for various* N_0, *measured with* $\tau_1 = 2$ min *and* $\tau_2 = 10$ min. *The error bars show the standard deviation of the collapsed curves using different values for* τ_1 *and* τ_2.

First, the basic idea of Fisher-like models that the interplay of cell motion and cell proliferation leads to a transition from a motility-induced spreading regime to a proliferation-dominated advancing front regime is well-verified. Then, the use of fast moving AX2 cells allowed us to quantify finer effects: we could notice the presence of secreted QSF through the decrease of cell motility, which is suppressed by continuous FM perfusion. Yet its macroscopic effect remains questionable. Above all, we could infer a previously unreported effect of local cell-cell interactions, which we name Contact Enhancement of Locomotion – or CEL – to contrast from the slowing down suggested by the term 'inhibition' in CIL (described in **Chapter 1**): the cells seem to increase their persistence upon cell-cell contacts. This last finding was confirmed by computer simulations, as described in the next section, and could play an important role in situations where cells spread out from an initial constrained location.

4 Modelling approaches

In this section, we propose various models, based on the concepts of the FKPP equation, to quantify the importance of the effects that we observed experimentally and to get a better insight in their mechanisms. To that end, we used both partial differential equations – PDE – models and individual-based ones.

Partial Differential Equations. The mean evolution of fields such as the cell density, some chemical concentrations, or other distributions, can be described by one or a set of Partial Differential Equations (PDEs). Most of the equations that we propose here are not easily analytically solvable, so we solved them numerically, using a Maple code on computers from Tohoku University, with the kind authorisation and help of Pr. Yoshinori Hayakawa.

1D cellular automaton. The particle-based models used here are of two types. First, I implemented a simple 1D cellular automaton, described in **Section 4.3** to get insight into the main mechanisms at stake.

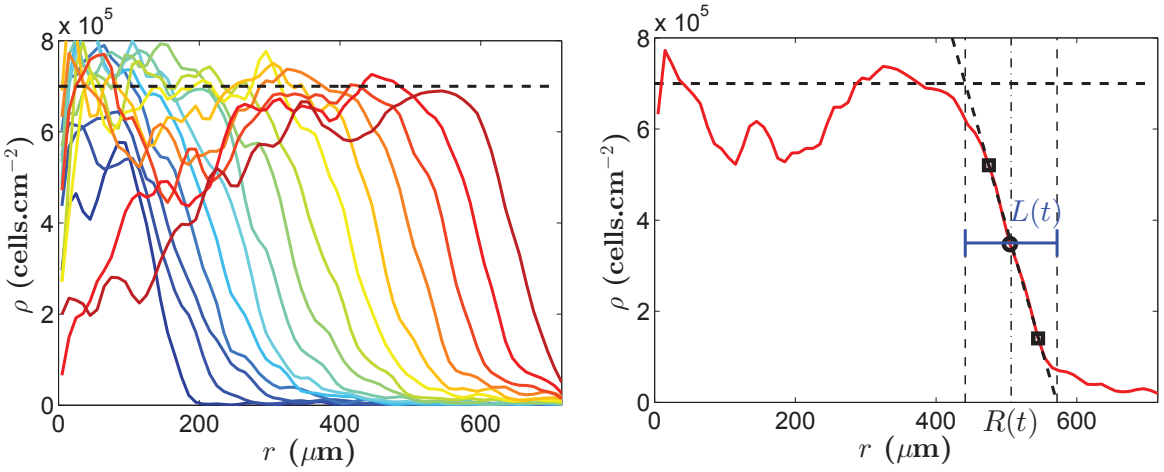
Active Brownian particles simulations. Second, we started a collaboration with Alexandre Solon (Physics Department, MIT, USA, and formerly MSC laboratory, Université Paris Diderot), to design 2D simulations of active particles to study the effect of various classes of contact interactions in the geometry of the experiments.

4.1 The limits of FKPP

It was clear even from $R_g(t)$ that FKPP would not be able to model the evolution of the AX2 colonies in any fashion. Therefore, we tried to compare the results of its numerical solving with those of DH1 colonies, which seemed to exhibit more constant and 'FKPP-like' dynamics only.

To that end, we chose to adopt an approach similar to that of Simpson *et al.* [110]: we take advantage of the bijection between $\{c, L\}$, respectively the asymptotic front speed and width, and the parameters $\{D, \lambda\}$. The former can be measured from the smoothed density profiles in two steps (Fig. 4.26b): first a rough, pixel-resolved estimate of $R(t)$ is made by detecting the first place where $\rho(r,t) < \rho_m/2$ (and $L(t)$ is estimated in the same fashion: $L(t) \simeq R_2(t) - R_1(t)$, where $\rho(R_1,t) = \frac{3}{4}\rho_m$ and

$\rho(R_2, t) = \frac{1}{4}\rho_m$); then these estimates are used to define the almost straight part of the profile, which we fit linearly to obtain $R(t)$ and $L(t)$ more accurately. L is computed as the mean of $L(t)$ for large t, and c is the slope of $R(t)$ on the same range of times.



(a) *(b)*

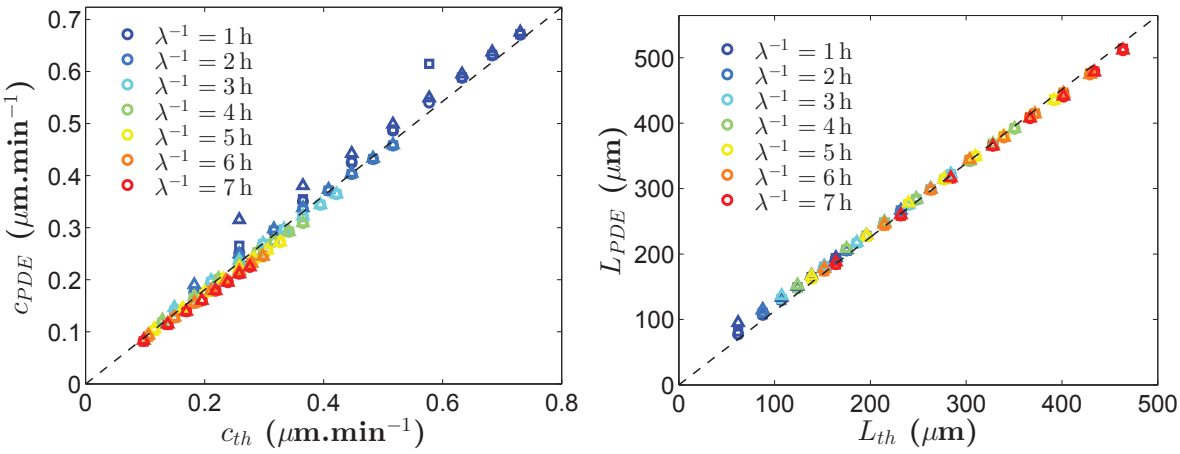
Figure 4.26: Wave fronts in density profiles.
(a) Density profiles from a typical spreading DH1 colony. $\Delta = 200$ min separation between subsequent curves. Note that the hollows at the back of late profiles are due to aggregates that form at high density (see ***Chapter 5****) and prevent the detection of cells. (b) Method for the measurement of $R(t)$ and $L(t)$: first a rough estimate of R is made (disk), and then the front (between the two squares) is fitted with a straight line. $R(t)$ is the point where this line equals $\rho_m/2$ (vertical dash-dotted line) while $L(t)$ is the length over which it decreases from ρ_m tp 0 (blue segment).*

We first tested by numerical solving of the FKPP equation (4.6) whether we retrieved the 1D predictions for c and L:

$$c = \sqrt{4D\lambda} \tag{4.17}$$

$$L = 8\sqrt{\frac{D}{\lambda}}. \tag{4.18}$$

Remember that $\lambda = \frac{\ln(2)}{T_2}$, where T_2 is the doubling time. For each set of parameters $[\lambda; D]$, we run three solvings, starting from the experimental initial conditions and using the corresponding carrying capacities ($\rho_m = \{5.5; 7; 7\} \times 10^5$ cells.cm^{-1} for experiments #1, #2 and #3 respectively). We find that expression 4.17 over-estimates slightly c, while expression 4.18 under-estimates L. Linear fits of the measured values to the predicted ones show that corrections by a factor 0.9 and 1.1 respectively yield accurate estimations (Fig. 4.27).



(a) *(b)*

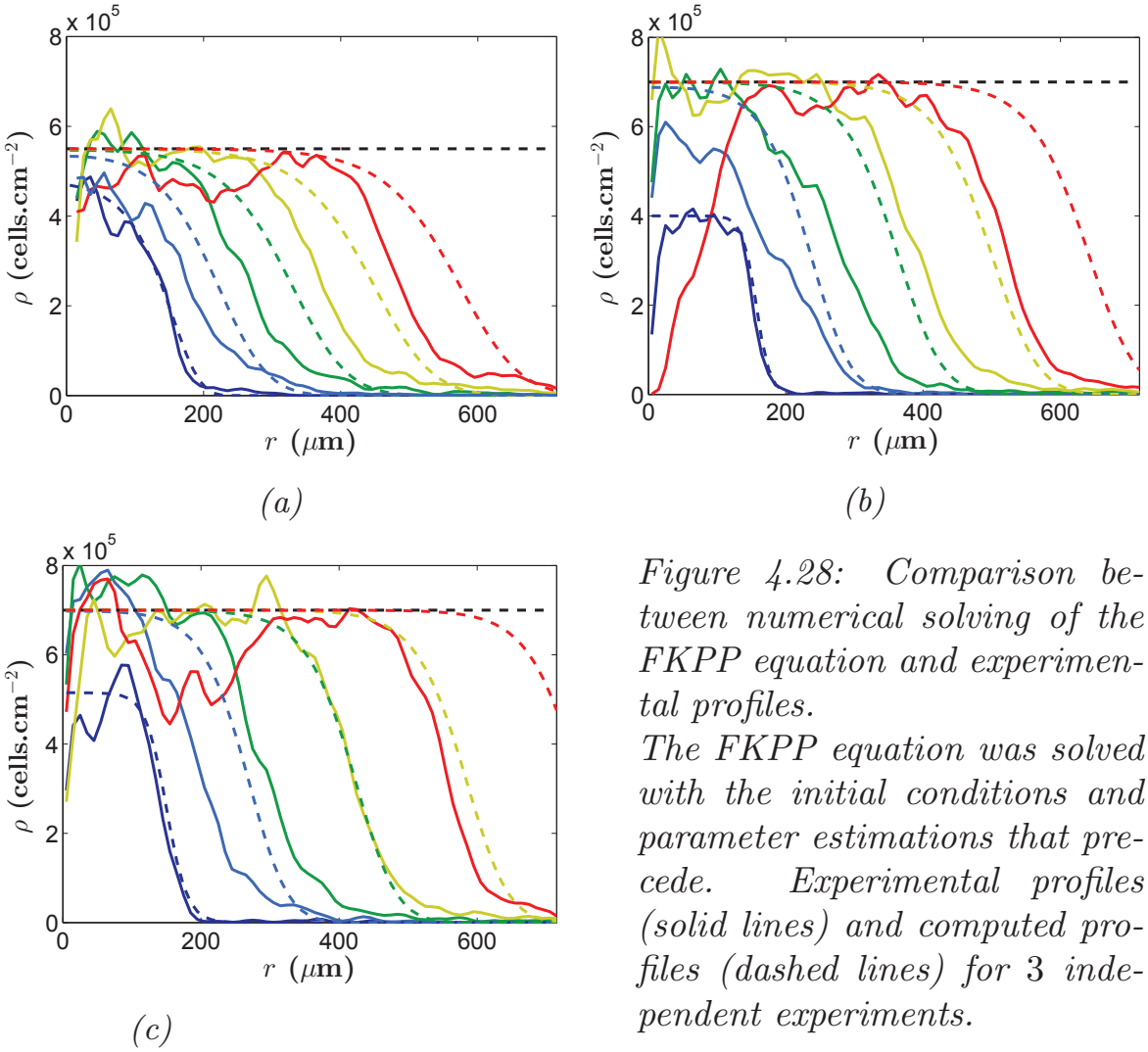
Figure 4.27: Prediction accuracy for the front speed c and width L. (a) c, measured from the profiles from numerical PDE solving as a function of theoretical values computed with 4.17, for various division times (λ^{-1} from 1 h to 7 h, from blue to red) and with parameters from 3 experiments (disks, squares and triangles). The slope of the dashed line is 0.9. (b) L measured from numerically obtained profiles as a function of the predicted values, from the same profiles. The slope of the dashed line is 1.1.

We then measured c and L in three independent experiments with DH1 cells (4.28, solid lines). From that we deduced the values of D and λ presented in Table 4.1. Then, we solved again the FKPP equation, starting from the experimental initial conditions, and with these estimated parameters (and with 10% variations in these parameters). The shape of the profiles is reproduced quite well, as well as the asymptotic front speed (Fig. 4.28, dashed lines). This is expected since this is how the parameters were optimised. However, the FKPP solution moves always ahead of the experimental front, with a lead of several hours (from 6 to more than 10 h!).

This effect is better understood by looking at the early shapes of $\rho(r,t)$: the experimental profiles seem to exhibit a retardation in proliferation with respect to the theoretical ones. Indeed, the profile does not spread out too much in comparison, but $\rho_{FKPP}(0,t)$ grows very fast towards ρ_m while it is not the case in experiments. Yet, we suspect this is mislead of a model-focused point of view. Indeed, the estimated doubling times (Table 4.1) are much shorter than the expected ones, around 9 h (Fig. 4.10). Conversely, the diffusion constant is smaller than that found by microscopic measurements. These discrepancies could reflect the fact that, as in the case of AX2, an outwards bias arises from cell-

Experiment	1	2	3	Expected
c (μm.min^{-1})	0.21	0.23	0.24	≈ 0.16
L (μm)	248	171	144	≈ 500
D (μm^2.min-1)	3.0	2.5	2.2	≈ 5
λ^{-1} (h)	3.7	2.6	2	13

Table 4.1: Measured values of c and L, and the deduced D and λ^{-1} from 3 DH1 experiments.



(a) *(b)*

(c)

Figure 4.28: Comparison between numerical solving of the FKPP equation and experimental profiles.
The FKPP equation was solved with the initial conditions and parameter estimations that precede. Experimental profiles (solid lines) and computed profiles (dashed lines) for 3 independent experiments.

cell interactions. (see 4.17). Such drift may lead to a sharper front [109, 111], which would allow higher motility and doubling time than the FKPP equation does.

Further work would be necessary to deeply understand these DH1 experiments: varying N_0 could tell more about the role of interactions for

instance. However, this example, with colonies that look at first sight to comply well with the FKPP equation, are a good demonstration of the limits of this model, and advocate the more thorough study of the effect of cell-cell interactions.

4.2 Downhill FKPP equation

Velocity down the density gradient. It seemed clear that modelling the motion with linear diffusion could not result in a good description of the experimental observation. In particular, two important features are not accounted for by it: the output of contact interactions, and the outward bias in the velocities – which can reasonably supposed to relate to each other. The last hypothesis was supported by a recent work by Bruna and Chapman [116], who studied the effect of hard-core repulsion in an assembly of (passive) Brownian disks. They showed that they resulted in an increase of the diffusion constant in dense zones. In particular, at high packing fraction, the diffusion constant writes (in 2D):

$$D = D_0(1 + 4\frac{\rho}{\rho_m}), \tag{4.19}$$

where $\frac{\rho}{\rho_m}$ represents the packing fraction. Similarly as in a previous work [109], this dependence of D on ρ can be interpreted as a drift velocity down the density gradient:

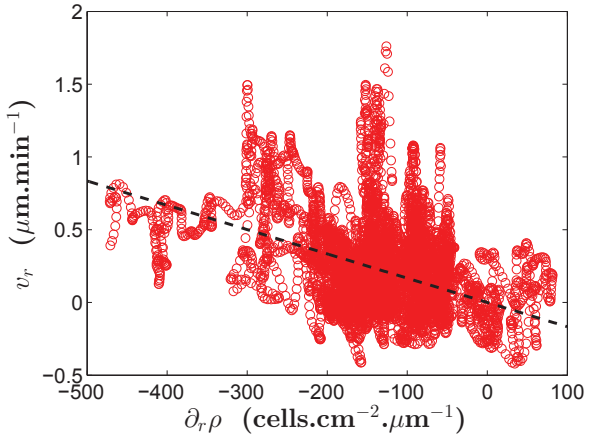
$$\begin{cases} \mathbf{j} = -D\boldsymbol{\nabla}\rho + \mathbf{v}_{drift}\rho \\ \mathbf{v}_{drift} = -\alpha\boldsymbol{\nabla}\rho \end{cases} \tag{4.20}$$

where $\alpha = \frac{4}{\rho_m}$ in the case of passive Brownian particles [116].

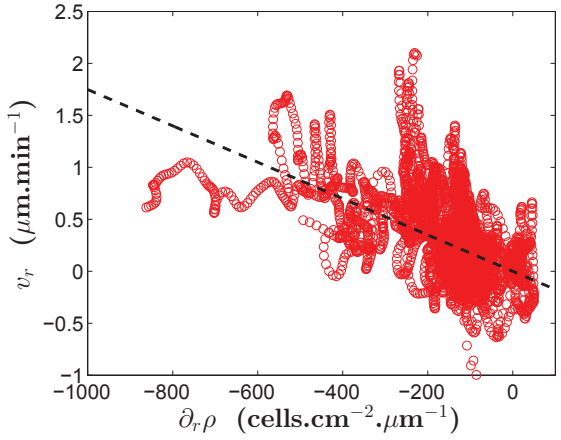
We plotted v_r versus $\boldsymbol{\nabla}\rho$ for all N_0 (Fig. 4.29). They seem to be linearly anti-correlated, *provided the points at $t < t^*$ are removed*, with a slope $-\alpha$ that depends on N_0. That means that the concept of velocity down the density gradient remains consistent with our data, up to two important features:

(i) it is true only after an initial rise of the polarisation;

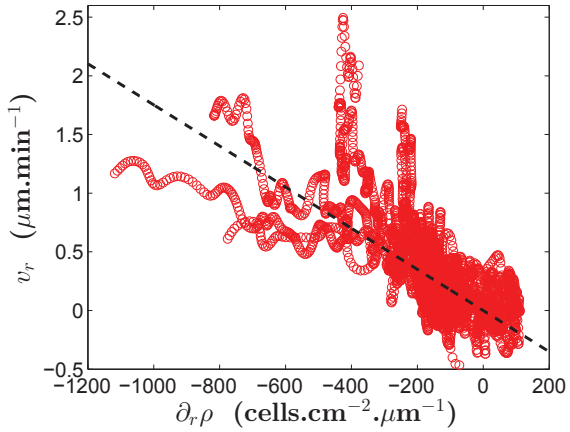
(ii) the relation between v_r and $\boldsymbol{\nabla}\rho$ is not an intrinsic property of the system but depends on the initial condition at $t = t^*$.



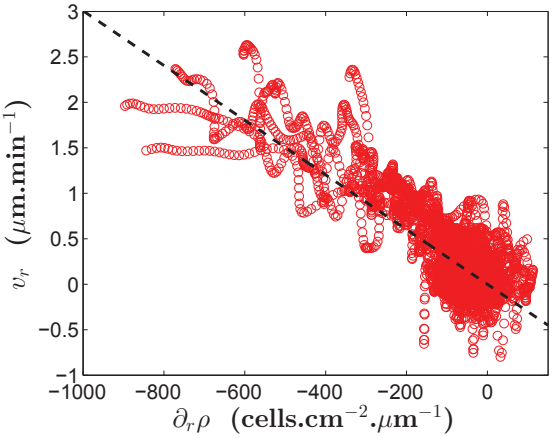
(a)



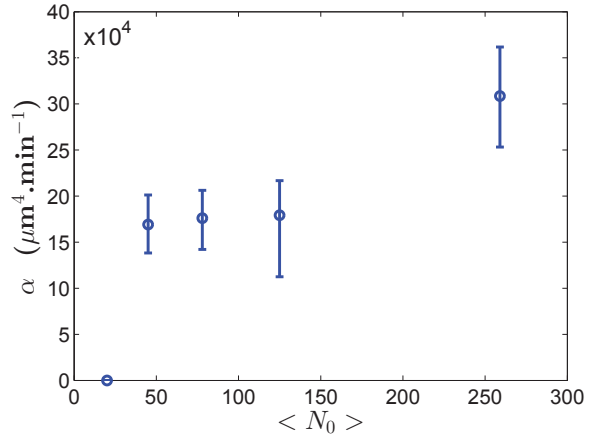
(b)



(c)



(d)



(e)

Figure 4.29: Velocity down the density gradient.
v_r versus $\partial_r\rho$ for (a) $\langle N_0\rangle = 45$, (b) $\langle N_0\rangle = 78$, (c) $\langle N_0\rangle = 125$ and (d) $\langle N_0\rangle = 259$. The dashed lines are linear fits $v_r = -\alpha.\partial_r\rho$. (e) $\alpha(\langle N_0\rangle)$. The error bars show the variations of the α estimate when the number of points in the (a–d) plots is changed.

If this drift was of the same origin as in Bruna & Chapman's case, α should be a constant for all N_0 and take the value

$$\alpha = 4\pi D_0 R^2 \sim 10^2 \ \mu\text{m}^4.\text{min}^{-1}, \tag{4.21}$$

for disks of radius $R = 5$ µm, which is much under the measured values $\alpha \sim 1 - 4 \times 10^5$ µm^4.min^{-1}. This means that the bias should originate from something else than the steric hindrance only. Nevertheless, we added such a term to the FKPP equation, to build what we term the downhill-FKPP equation:

$$\partial_t \rho = \boldsymbol{\nabla} \left((D + \alpha\rho) \boldsymbol{\nabla} \rho \right) + \lambda \rho \left(1 - \frac{\rho}{\rho_m} \right). \tag{4.22}$$

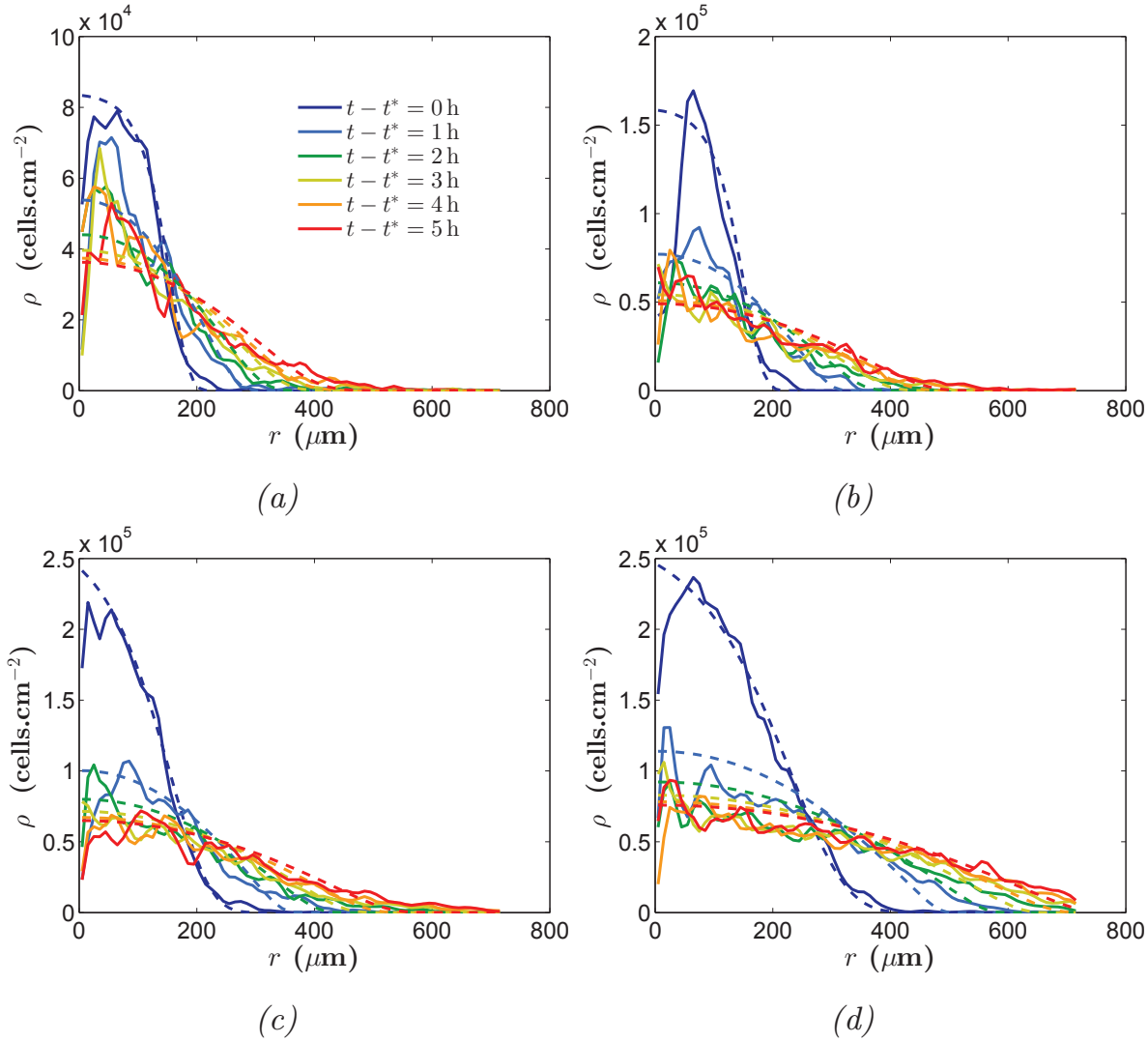


Figure 4.30: Downhill-FKPP results.
Density profiles from experiments (solid lines) and numerical solving of the downhill-FKPP equation (dashed lines). (a) $\langle N_0 \rangle = 45$, *(b)* $\langle N_0 \rangle = 78$, *(c)* $\langle N_0 \rangle = 125$, *(d)* $\langle N_0 \rangle = 259$. *The parameters are fixed to* $\lambda^{-1} = 9$ h, $D = 10$ µm^2.min^{-1}, $\alpha = \alpha(N_0)$ *from Fig. 4.29e.*

Taking the density profile $\rho(r, t^*)$ for each N_0 as the initial condition, we solved numerically this equation over 500 min. λ was fixed to 9 h consistently with the data, and we used the measured values of $\alpha(N_0)$.

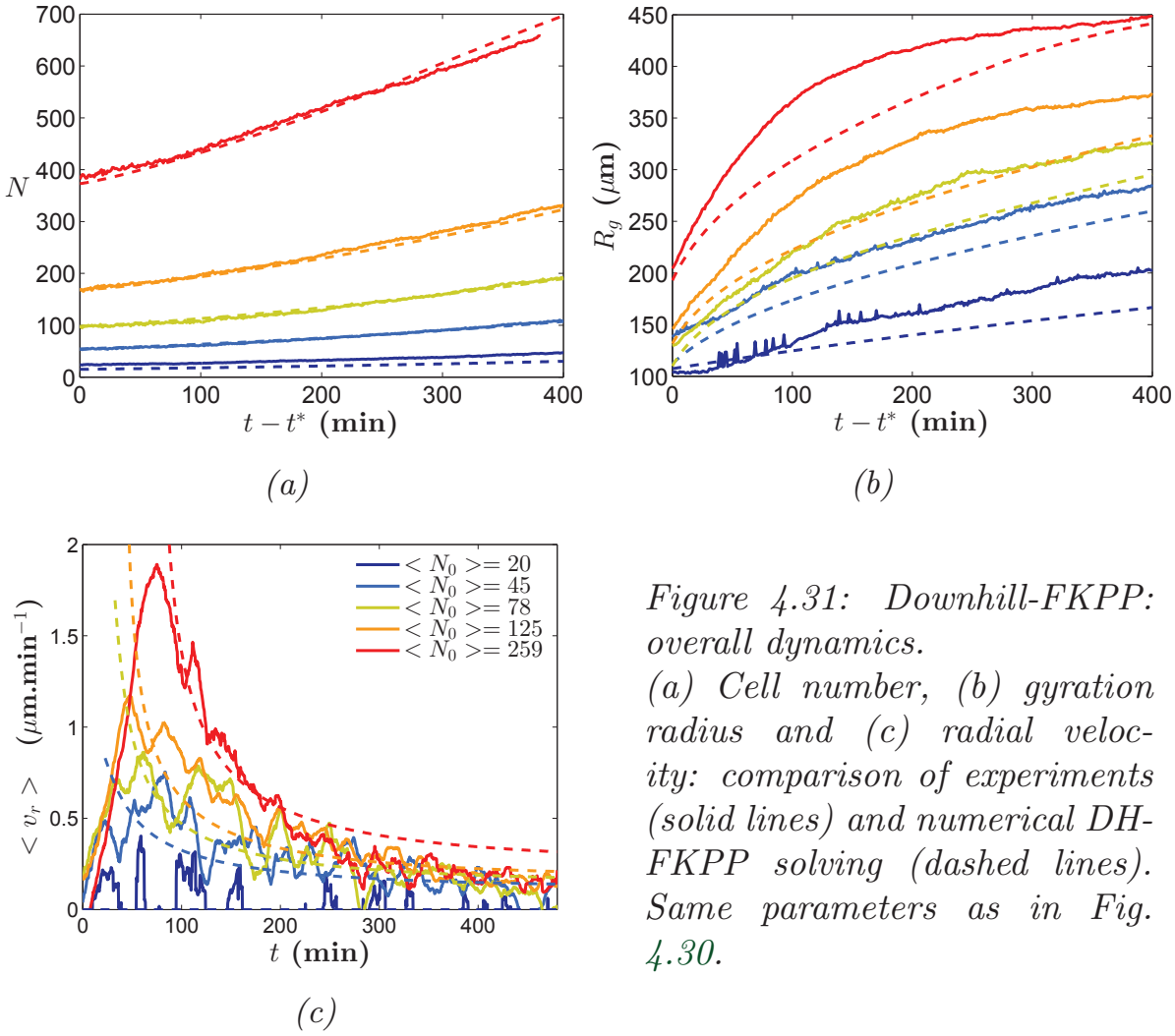
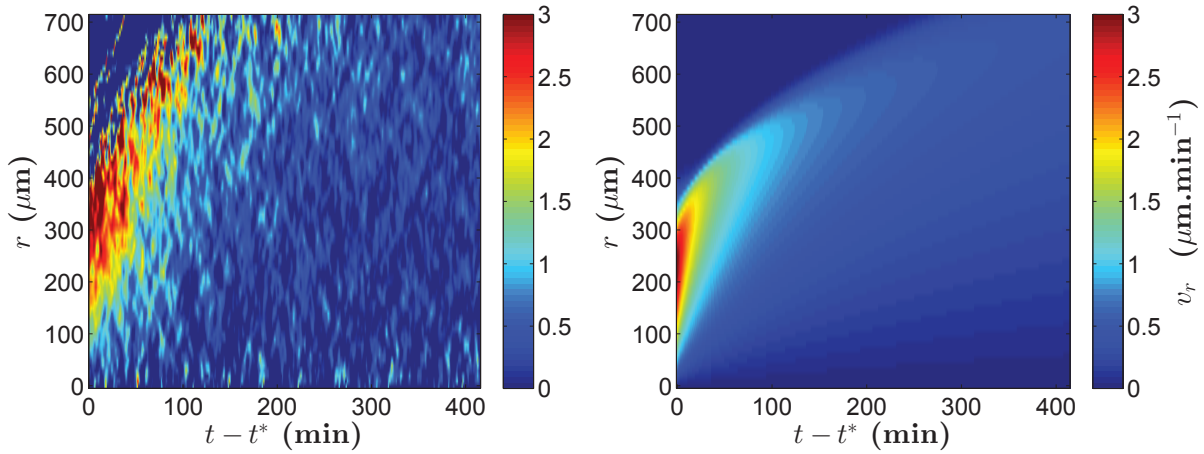


Figure 4.31: Downhill-FKPP: overall dynamics.
(a) Cell number, (b) gyration radius and (c) radial velocity: comparison of experiments (solid lines) and numerical DH-FKPP solving (dashed lines). Same parameters as in Fig. 4.30.

The most problematic parameter to set was D. For instance, the MSD is affected by the drift velocity, hence it cannot give D directly. However, we could estimate that $D = 10\ \mu m^2.min^{-1}$ is a good approximation for – and actually we checked that the dynamics is not highly affected by its value provided it is in the range $5 - 40\ \mu m^2.min^{-1}$.

Rather than allowing the parameters to move in order to fit the data, we decided to check the behaviour of the downhill-FKPP equation with the parameter values inferred from the microscopic measurements. The match of the computed profiles and of various macroscopic quantities to the data is displayed on Figures 4.30 and 4.31 respectively. It shows that this equation captures the essential dynamics of the evolution after t^*. In particular, even for times smaller but comparable to the doubling time, the growth term is necessary to match the profiles and $N(t)$ well. Indeed, in its absence the profiles spread too fast. The decay of $\langle v_r \rangle(t)$ is also quite well described, even though it is much over-estimated in the first minutes of the solving, for $\langle N_0 \rangle \geq 78$. Note that the data could be



(a) *(b)*

Figure 4.32: Comparison of the spatio-temporal dynamics of v_r in (a) experiment and (b) numerical solution of the downhill-FKPP equation, starting from the same initial condition ($\langle N_0 \rangle = 259$)) and parameters as in Fig. 4.30.

better fitted by changing the parameters' values, but we did not carry out such a procedure as it would not have been very informative.

The spatio-temporal dynamics are also quite consistent with the experimental ones (Fig. 4.32): $v_r(r,t)$ is high on the edge of the colony, where the density gradient is high, and it relaxes as the profiles become shallower. By contrast, this equation is not able to produce the initial polarisation phase: rise of v_r from 0 and subsequent peak at $t = t^*$. Starting from a very steep gradient concentrated at the very edge of the colony, it is expected that this gradient propagates inwards as the density spreads out. Yet, this phenomenon happens on a very short time-scale of the order of a few minutes and so it cannot account for the slow signal propagation that we observe experimentally.

To sum up, the results of the downhill-FKPP equation confirm that an outwards drift velocity, concentrated in the zones of high density gradient, is the driving mechanism of the colony spreading at $t \geq t^*$. However, it also confirms that the magnitude of this drift depends on its initial state at t^*, but it is unable to explain the dynamics before t^*.

4.3 Polarisation and persistence: insights from 1D simulations.

Once it is evidenced that the colony dynamics is controlled by the polarisation, and that the polarisation is driven by contact interactions,

the interaction mechanism at play remains to be elucidated – and the passive hard-core repulsion is seemingly not enough. The approach of cell biology would be to screen various mutants, looking for a change in the response to these interactions, together with a thorough analysis of intracellular processes in the contact zone, for instance by fluorescence microscopy. We chose to tackle this problem with the view of physics, aiming at identifying the minimal class of interaction which can give rise to the observed collective phenomenon.

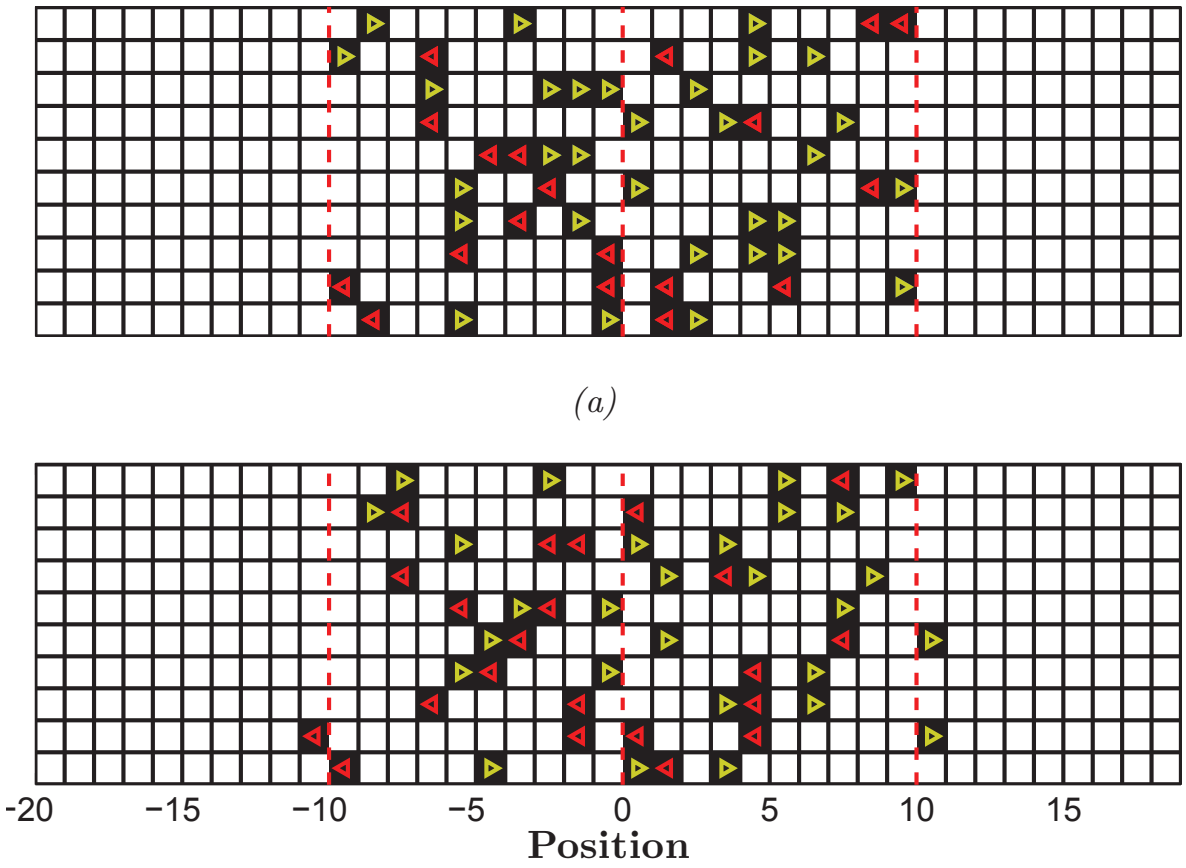


Figure 4.33: Principle of the 1D cellular automaton.
(a) At $t = 0$, $N = 5$ cells are distributed at $-L_i/2 < r < L_i/2$. Here $L_i = 20$, the borders materialised by the two vertical red lines. 10 independent simulations are shown. Then they are attributed a polarity, illustrated by the arrowheads ("+1": green, to the right, "-1": red, to the left). (b) At $t = 1$, the cells that were free to move have moved to their target site; those facing a contact have tumbled without moving; those 'fighting' for a target have either moved or not (example: bottom line, the cell at -0.5 has moved to the right, while that at 1.5 has not moved).

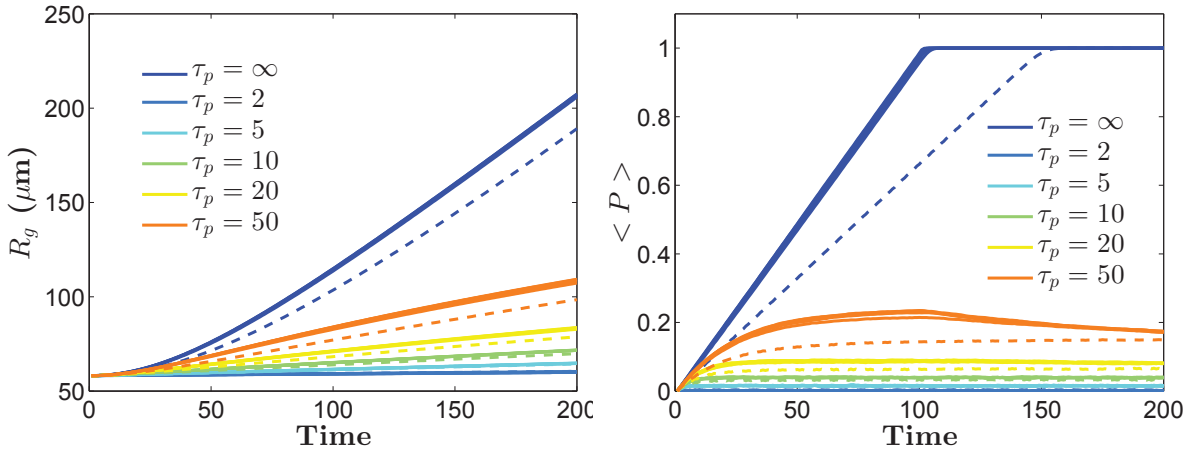
A one-dimensional (1D) cellular automaton provides a convenient tool to

that end, since what happens in the system is very easily interpretable. It was built as follows (see Fig. 4.33):

- the space is defined as a line with equally spaced sites (squares in Fig. 4.33);
- one site can be occupied either by zero (white) or one (black) cell;
- the cells have a $+1$ (green arrow-head) or -1 (red arrow-head) polarity;
- at each time step all cells i move onto the neighbouring site in the direction defined by its polarity p_i, then they can tumble (reverse their polarity) with a rate $\frac{1}{\tau_p}$;
- if the neighbouring site is not free, the cell does not move and tumbles;
- if the neighbouring site is free but a cell coming from the opposite side aims at it as well, one of the two is chosen at random and does not move, while the other is allowed to move to its target.

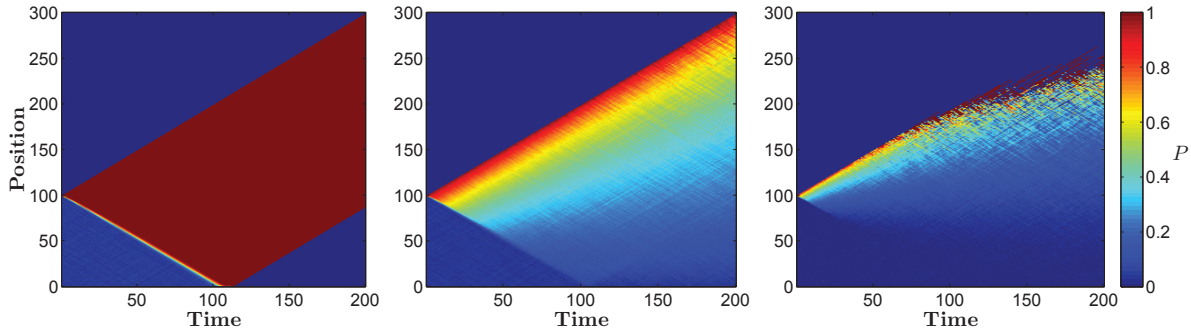
To model circular colonies in 1D, we used a segment of length L_i randomly occupied by N cells of random polarity as the initial condition and then let the system evolve following the above rules.

Note that in its first basic version, this automaton has symmetric collisions, to the extent that a cell has the same properties before and after a collision. Starting with N 'cells' concentrated on a segment of width $L_i = 200$, only τ_p plays a role in the overall dynamics (Fig. 4.34), except for $N \sim L_i$ due to the initial crowding that slows down the system (when two cells aim at the same target, only one moves). The cell 'size' does not change anything either. To check that, we multiplied all lengths and times – L_i, τ_p, observation time and length – by a factor 2 or 5, which is equivalent to a decrease of the spacing between sites by the same factor: keeping the same number of cells, there are more free sites between adjacent cells. The dynamics remained unchanged in these cases (not shown).



(a) *(b)*

Figure 4.34: Overall dynamics of the 1D automaton.
(a) Gyration radius and (b) polarity (analogue to $\langle v_r \rangle$) for 'cell colonies' simulated with the 1D cellular automaton with various parameters. The persistence time τ_p is colour-coded, $N \in \{5; 10; 20; 100\}$ (full, superimposed lines) and $N = 200$ (dashed lines). $L_i = 200$. $10,000$ independent simulations.



(a) *(b)* *(c)*

Figure 4.35: 1D automaton: spatio-temporal dynamics of the polarity. Polarity maps for (a) $\tau_p = \infty$, (b) $\tau_p = 50$, (c) $\tau_p = 10$. $N = 100$. $L_i = 200$. $10,000$ independent experiments.

Polarity due to persistence. Looking at the polarity $P(r,t) = \langle p_i \rangle$, where the average is taken over the two sites at a distance r of the center in 10.000 simulations, one observes a wave that propagates from the edge to the center at the speed of the cells, $v = 1$ (Fig. 4.35).

These results can be interpreted as follows. Since the collisions are symmetric, this system is actually similar to a system of particles moving with a persistence τ_p *with no interaction*. Indeed, when a couple of cells collide, the output is the same as that of a crossing, provided the cells' are not labelled. This explains why N has no effect.

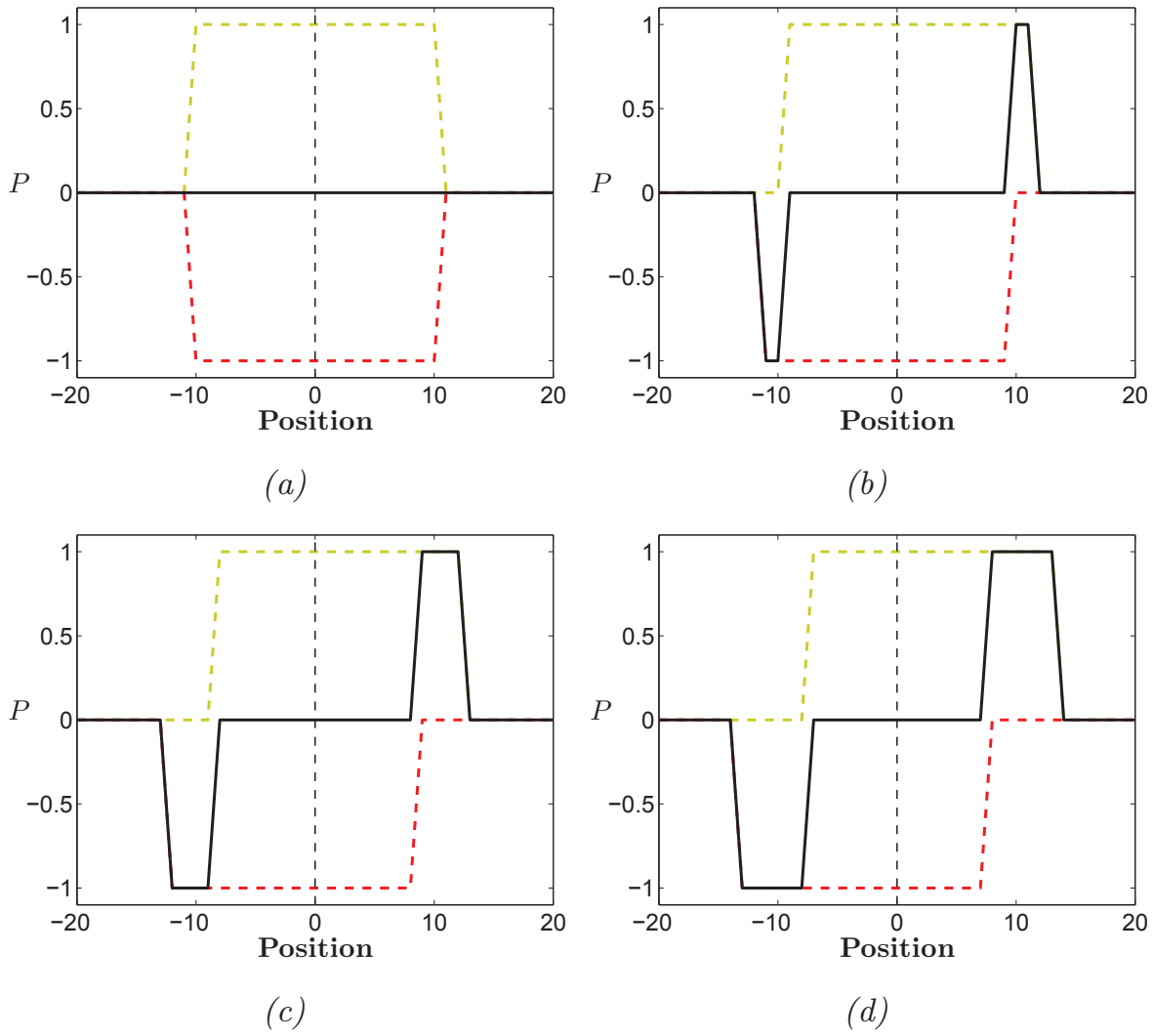
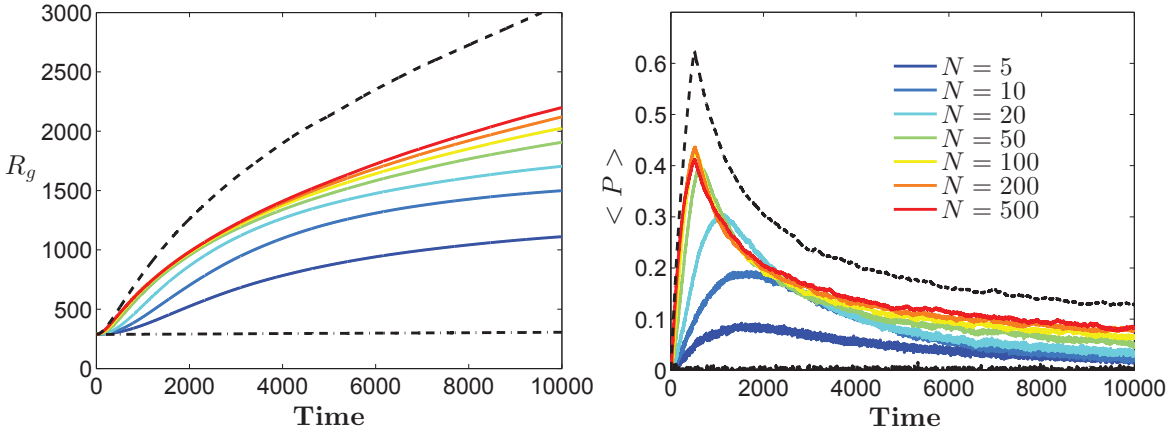


Figure 4.36: Origin of the polarity without interactions. Distribution of 'cells' with infinite persistence. The dashed lines show the density distribution of the +1 (green), and the -1 cells (red), while the black solid line represents the local average polarity. At $t = 0$ (a), the mean polarity is 0 everywhere, but as time goes both population separate, inducing the growth of polar zones on the edges (b, c, d: $t = 1 - 3$).

Thus, the polarity wave comes only from the persistence. To illustrate this, think of non-interacting cells, whose polarity is initially set at random in the colony, with *infinite* persistence (as schematised in Fig. 4.36). The $N/2$ "-1" cells will move as a solid to the left, while the $N/2$ "+1" cells will move to the right. That creates a band that contains only "-1" on the left edge (and symmetric) and that grows with speed $v = 1$ in both directions until it reaches the center. Then the groups will separate and move away from each other.

Now, if the cells have a finite persistence, for $t \ll \tau_p$ the system behaves just as described, while at $t \gg \tau_p$ the cell density tends to be Gaussian

and the average polarity remains very low. The polarity wave with speed $v = 1$ still builds up, but it is counterbalanced by the tumbles on a time $\sim \tau_p$. This is why it disappears on long time scales, leading to a peak of average $\langle P(t) \rangle_r$.



(a) *(b)*

Figure 4.37: 1D simulation of Contact Enhancement of Locomotion (a) Gyration radius and (b) polarity for simulated 'cell colonies' with $L_i = 1000$, $\tau_p = 2$, $\tau_e = 1000$, $N \in \{5; 10; 20; 50; 100; 200; 500\}$ *(colour-coded solid lines). The limit cases with no CEL are plotted in black with* $\tau_p = 2$ *(dash-dotted lines) and* $\tau_p = 1000$ *(dashed lines).*

Asymmetric collisions: Contact Enhancement of Locomotion. In this particular system, the only way to produce a collective effect is to consider asymmetric collisions: as a collision is undistinguishable from a crossing, the state of the system will be changed upon contact only if the state of the cells themselves is changed: for instance, the cells can either escape with a greater speed or a greater persistence. Since the frequency of collisions increases with density, the spreading should increase with N.

The second option was implemented as follows: after a collision, the persistence of both colliding cells is increased to $\tau_e > \tau_p$. Then it relaxes to τ_p at the next tumble, or it is set again at τ_e at the next collision.

To explore a broader ranger of persistence, we increased the initial segment length L_i to 1000 and the observation time accordingly (this process is equivalent to decreasing the particle size). Now there is a clear effect of the cell number (Fig. 4.37). In particular, at high (*resp.* low) N the behaviour approaches that of a system with symmetric collisions and persistence τ_e (*resp.* τ_p).

4.4 Active Brownian particles: 2D simulations.

NB: the work reported in this section is based on interacting particles codes written and implemented by Alexandre Solon.

We used 2D simulations, based on the active Brownian particle model, to get a more precise idea of the effect of contact interactions in two dimensions and in continous space.

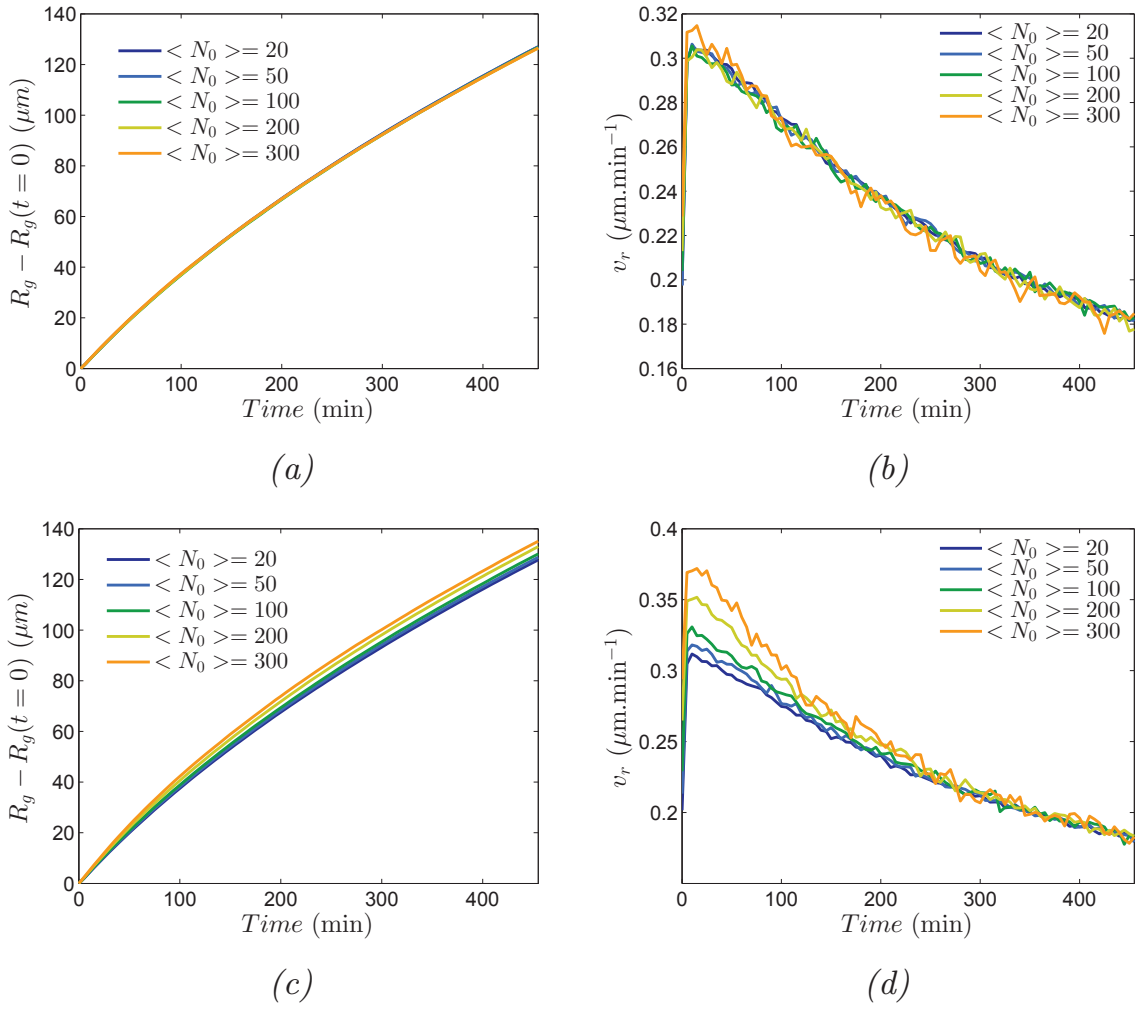


Figure 4.38: Active Brownian particles with simple collision rules. (a,c) Gyration radius and (b,d) radial velocity for colonies of active Brownian particles undergoing simple interaction rules: hard-core repulsion only (a,b) or hard-core repulsion and angular repulsion (c,d). The particle were initially constrained in a disk of diameter 320 µm. $v = 5$ µm.min^{-1}, $D_r^{-1} = 5$ min, *(a,b)* $\gamma = 0$ *or (c,d)* $\gamma = 100$.

Hard-core repulsion only. We started with steric hindrance only, modelled by a hard-core repulsive force acting on touching particles. The dynamics of particle i is described by the following set of stochastic equations:

$$\partial_t \mathbf{r}_i = v\mathbf{u}(\theta_i) + \sum_{j\neq i} \mathbf{f}_{ij}(\mathbf{r}_i - \mathbf{r}_j) \tag{4.23}$$

$$\partial_t \theta_i = \sqrt{2D_r}\eta_i(t) \tag{4.24}$$

where $\mathbf{u}(\theta_i) = (\cos\theta_i, \sin\theta_i)$ and η_i is a delta-correlated Gaussian white noise with zero mean and unit variance. $\mathbf{f}_{ij}$ is the steric repelling force exerted by particle j on particle i, and derives from a Weeks-Chandler-Andersen potential $V(r) = 4\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + 1$ if $r < 2^{1/6}\sigma$ and 0 otherwise, where $\sigma = 10\,\mu\text{m}$ is the particle radius. The Langevin equations Eqs. (4.23-4.24) were integrated using a Euler integration scheme with time steps $\Delta t = 10^{-3}$ min.

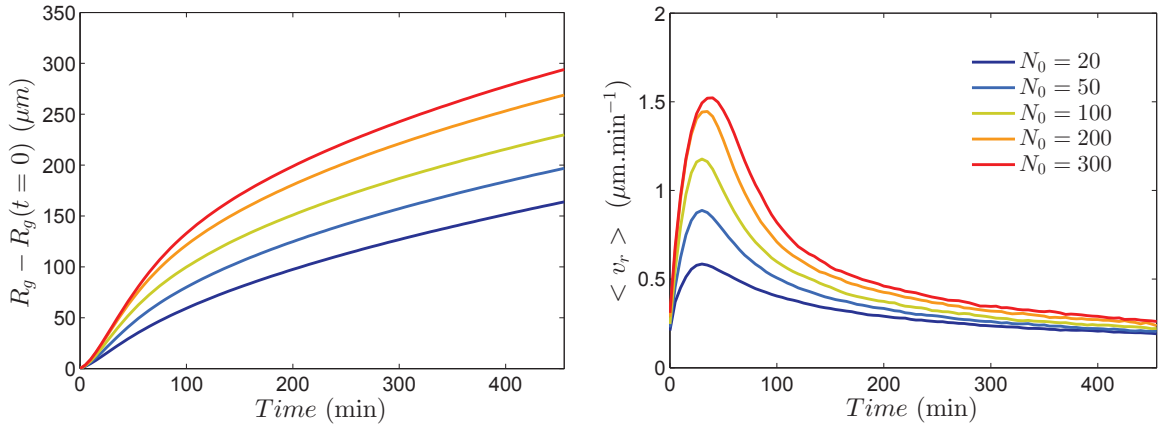
As we have seen in **Section 4.2**, this effect alone can increase the diffusion constant in a density-dependent way. It is due to an additional pressure arising from interactions, which has also been described for active particles [117]. However, this effect is expected to be significant only at close packing, which is not reached in experimental conditions (packing fraction up to ~ 0.3, assuming that the effective interaction radius is the cell radius itself), and it should even be counterbalanced by the decrease of the effective speed in the case of ABPs [55]. Indeed, there is no effect of the particle number on the dynamics (4.38a-b).

Active reorientation through angular repulsion. Contrary to the 1D automaton, the collisions do not induce a change of direction in the latter model. Yet, it is reasonable to think that the cells actively reorient upon contact: this is the current paradigm of the contact inhibition of locomotion (CIL), that the protrusions are inhibited in the contact zone, leading to a repolarisation in the direction away from the contact. This was implemented by adding a angular repulsion between two interacting – *ie* at contact – particles:

$$\begin{cases} \partial_t \mathbf{r}_i = v\mathbf{u}(\theta_i) + \sum_{j\neq i} \mathbf{f}_{ij}(\mathbf{r}_i - \mathbf{r}_j) \\ \partial_t \theta_i = \sqrt{2D_r}\eta_i(t) + \gamma \sum_{j\neq i} \sin(\theta_i - \beta_{ij}) \end{cases} \tag{4.25}$$

where the term $\gamma \sum_{j\neq i} \sin(\theta_i - \beta_{ij})$ is a repulsive torque of magnitude γ, which reorients the particles away from their neighbours, with $\beta_{ij} = \arg(\mathbf{r}_j - \mathbf{r}_i)$.

Qualitatively, it could act to reorient preferentially the particles in the direction of negative density gradient. Indeed, there is a small increase in both v_r and the spreading rate when the particle number is increased (Fig. 4.38c-d). Yet, this increase remains very small and the peak is still very early, showing that angular repulsion is not sufficient to account for the spectacular effect observed experimentally (Fig. 4.18).



(a) *(b)*

Figure 4.39: Active Brownian particles undergoing CEL.
(a) Gyration radius (variation) and (b) radial velocity for colonies of active Brownian particles undergoing contact enhancement of locomotion: upon collision, they enter a ballistic mode that relaxes to normal persistent random motion with rate ω. The particle were initially constrained in a disk of diameter 320 µm. $v = 5$ µm.min^{-1}, $D_r^{-1} = 2$ min, $\omega^{-1} = 10$ min.

Contact Enhancement of Locomotion. Therefore, we implemented the effect of Contact Enhancement of Locomotion in a similar way as in the 1D automaton. Upon collision, the particles start a ballistic run until they relax stochastically to their natural mode, at a rate ω^{-1}. Here, we take $D_r^{-1} = 2$ min and $\omega^{-1} = 10$ min, corresponding to relaxation times close to τ_1 and τ_2 respectively, as measured in the experiments, and $\gamma = 0$ (no angular repulsion). Alternatively, the second mode can be an ABP mode with a higher persistence τ_e, and the speed in this mode can be changed – increased – as well.

The simulation results demonstrate that this effect alone (the angular repulsion has been switched off there) captures very well the main features of the experiments (Fig. 4.39). In particular, the quantitative agreement, in terms of values of $\langle v_r \rangle$ for instance, is surprisingly good given the simplicity of the model: here $v_r(t^*) \approx 1.5$ µm.min^{-1} for $N_0 = 200$,

while experimentally $v_r \approx 2$ µm.min^{-1} for $\langle N_0 \rangle = 259$ (Fig. 4.18d).

The main discrepancy is in the peak timing t^*, which is earlier in the simulations and does not seem to depend much on N_0, the particle number. Again, it seems to be set by the persistence time, namely ω^{-1} in this case, with only a change in proportion with N_0 (Fig. 4.40).

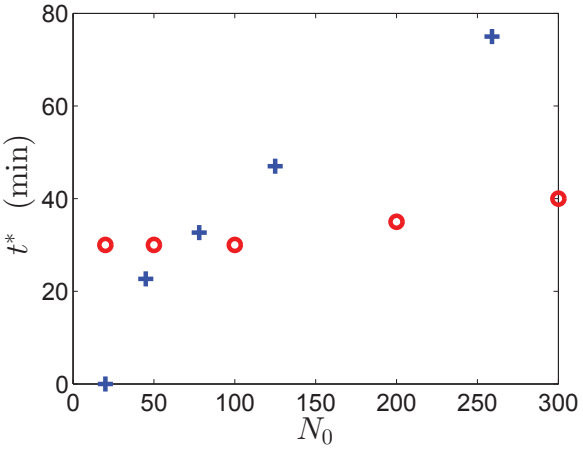


Figure 4.40: Timing t^* of the peak in radial velocity.
t^* as a function of N_0 from the experimental data (blue plus signs) and the bimodal ABP model (red circles).

The velocity autocorrelation function is well described by a sum of two exponential decays, as expected (François Detcheverry, personal communication, see **Appendix A**). Here, unlike in the experiments, $\phi_1 + \phi_2 = 1$, because only two modes are considered, with no positional noise. Fitting ϕ_2 and τ_2 from the expression

$$C(\delta t) = (1 - \phi_2) \exp(-\frac{t}{2}) + \phi_2 \exp(-\frac{t}{\tau_2}) \tag{4.26}$$

we find that ϕ_2 correlates well with the actual proportion of cells in mode 2. Furthermore, τ_2 increases as well in the denser, more persistent zones (not shown). This is consistent with a shorter mean free path, hence a shorter characteristic time for the duration of "random" runs, in mode 1.

4.5 Summary

In this section, we introduced a few concepts that allowed us to draw a more precise picture of the spreading dynamics of our *Dictyostelium* model colonies. First, by numerical solving of the FKPP equation, we demonstrated that this equation could describe well the long-time, proliferation-driven evolution of DH1 colonies, at the cost of a small distortion of the parameters; however, even in this seemingly FKPP-friendly case, it failed to account for the early dynamics.

This is apparently due to a bias in the motion, which is directed outwards in the early stage of the spreading. This bias, more visible in AX2 colonies on which all the following was based, can be relatively well introduced in the form of a drift velocity down the density gradient, for $t \geqslant t^*$. This gives birth to the downhill-FKPP equation, which predicts quite accurately the dynamics in this time frame ($t^* \leqslant t \leqslant 500$ min).

Yet, this phenomenological approach does not inform about the origin of the bias, and especially about its emergence at $t < t^*$. The latter can be seen as a polarisation induced by the cell persistence. Particle-based simulations showed that in order for it to exhibit the experimentally observed characteristics, the cells should undergo CEL, a process which increases their persistence upon cell-cell contact.

5 Accumulation model: getting closer to the data

Using particle-based simulations, we could quite satisfactorily explain the appearance of transient polarisation that relies on a simple mechanism of Contact Enhancement of Locomotion. However the precise dynamics of the polarisation is not completely captured by this simple two-state model. In particular, it does not exhibit the right $t^*(N_0)$ dependency (see Fig. 4.40) and it does not reflect the experimental observation that before t^* all the $\langle v_r \rangle(t)$ curves seem to be undistinguishable. Here we develop qualitative arguments to build a model that fits this behaviour, from the dynamics to the mechanisms.

We simplify the problem as follows: we consider an expanding domain, whose expansion dynamics is controlled by an expansion speed V, which is in a way the analogue of v_r in this model. Then we suppose that V is accumulated by contact interactions, so in a density dependent way, and is destroyed at a constant rate. The dynamics is given by:

$$\begin{cases} \dot{R} = V \\ \dot{V} = a(\rho) - \beta V \\ \rho = \frac{N}{\pi R^2} \end{cases} \tag{4.27}$$

The driving idea is that the peak in V is set by a transition from an accumulating regime ($a(\rho) > \beta V$) to a decay regime ($a(\rho) < \beta V$). The

critical functional is thus the accumulation rate $a(\rho)$, which can set t^* since ρ decreases as the colony spreads out (Fig. 4.41). The simplest shape possible for $a(\rho)$ is a stepwise function:

$$a(\rho) = \begin{cases} \beta V_0 \text{ if } \rho > \rho^* \\ 0 \text{ otherwise} \end{cases} \tag{4.28}$$

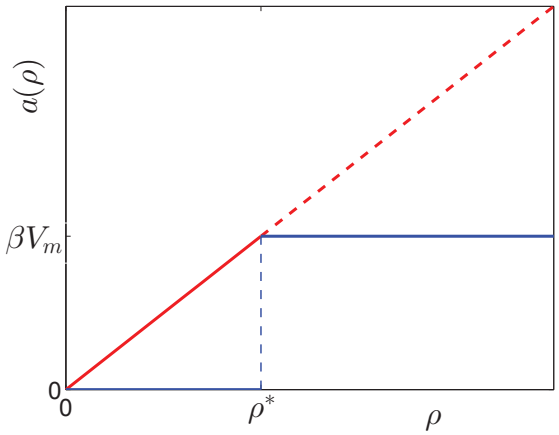


Figure 4.41: Possible shapes for the accumulation rate $a(\rho)$. In the sharp version (blue), a transitions instantaneously from βV_m to 0 when ρ passes down ρ^*. In the smooth version (red), $a \propto \rho$, with a potential saturation at βV_m, leading to a slow decrease of V following ρ.

With this assumption, the dynamics of V is very simple: as long as $\rho > \rho^*$, $V = V_m(1 - \exp(-\beta t))$, and at $t \geqslant t^*$ so that $\rho(t^*) = \rho^*$, $V = V^* \exp(-\beta(t - t^*))$. Importantly, one notices directly that for all N, as long as $V(t)$ increases all the curves collapse on the same one. Then t^* is determined by the condition $\rho(t^*) = \rho^*$.

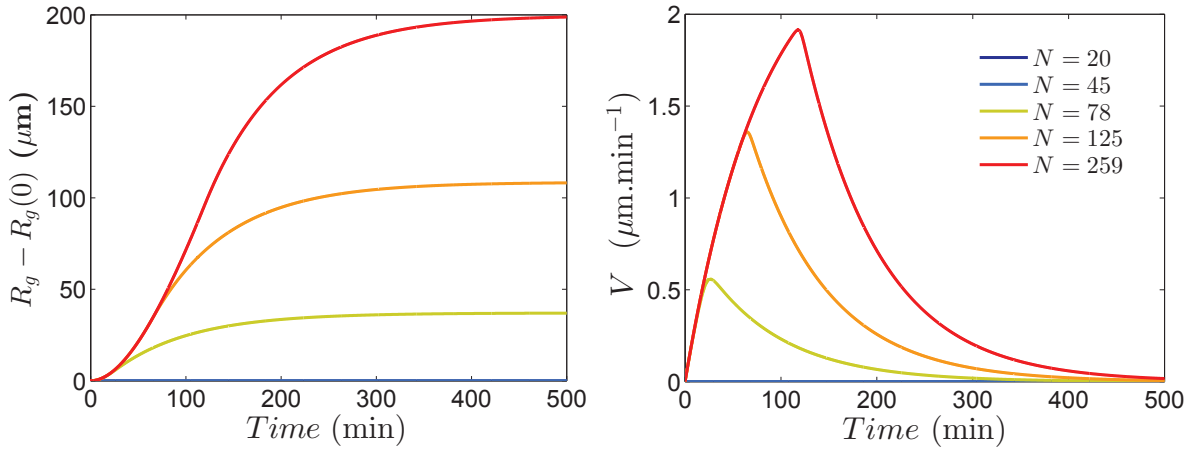
For $t \leqslant t^*$, the radius writes

$$\begin{aligned} R(t) &= R_0 + \int_{t'=0}^{t} V(t')dt' \\ &= R_0 + V_m t + \frac{V_m}{\beta}\left(e^{-\beta t} - 1\right), \end{aligned} \tag{4.29}$$

and so the condition on t^* yields

$$\begin{aligned} &\frac{N}{2\pi R(t^*)^2} = \rho^* \\ \Leftrightarrow &V_m t^* + \frac{V_m}{\beta} e^{-\beta t^*} = \sqrt{\frac{N}{2\pi\rho^*}}. \end{aligned} \tag{4.30}$$

The scaling of t^* with N is not obvious and depends on the relative values of the three parameters. We solved this model using a finite difference method with explicit scheme for a few sets of parameters. We observe the expected behaviour: an initial increase that is shared by all conditions N, and then an exponential relaxation from $t^*(N)$ (Fig. 4.42b). Moreover, t^* increases with N_0 in a way that reminds the experimental data (see Fig. 4.44, dark green triangles).



(a) *(b)*

Figure 4.42: Accumulation model – sharp version.
(a) $R_g = \frac{2}{3}R$ *and (b)* V *as a function of time obtained from the accumulation model.* $a(\rho \geqslant \rho^*) = \beta V_m$, $a(\rho < \rho^*) = 0$, *with* $\beta^{-1} = 80$ min, $V_m = 2.5$ µm.min^{-1}, $\rho^* = 9 \times 4$ cm^{-2}.

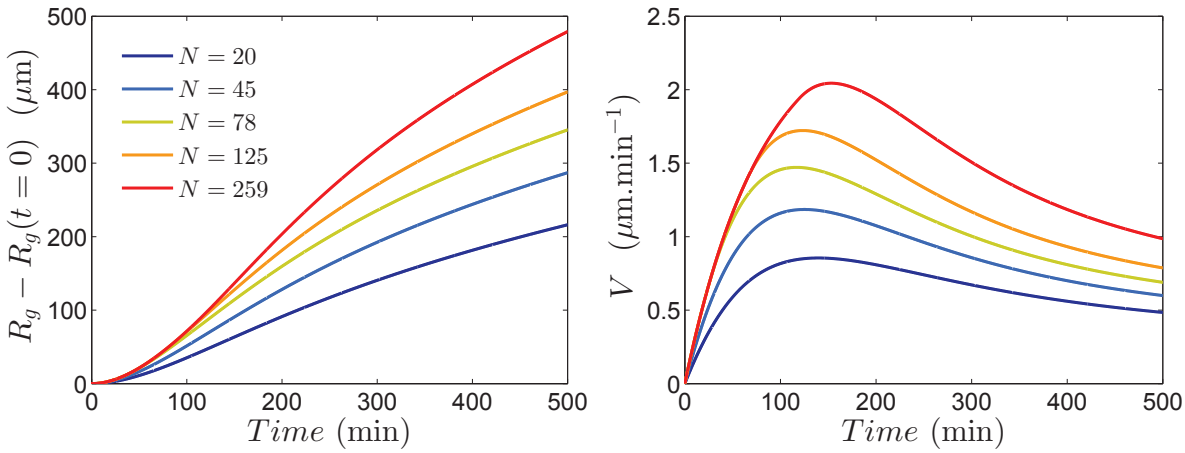
Of course, there are discrepancies with the experimental observations. As R is only driven by V, it tends to a constant. This is because in this simple avatar, our model does not account for any spreading in the absence of V, which represents the polarised motion. It could be overcome by including an additional term in $\dot{R}$, for instance in the form of a basal speed V_0, or even of a diffusion constant (so that $\dot{R}^2 = 4D + 2V^2t$ for instance); alternatively $a(\rho)$ could be shaped so that V does not relaxes to 0.

As such, it is only a toy model that allows to estimate the behaviour of a colony where persistence for instance – or speed – would be accumulated upon contact, instead of simply switching from a native state to an activated state. It shows that this hypothesis has some potential to better explain the experimental data and could deserve being integrated in a more realistic particle-based or PDE model.

Biologically speaking, it could rely on the production of internal polarising or activating signals by the cells upon cell-cell contact, which signals

would be degraded at a constant rate β. If a given amount of this signal was produced *per* collision, then $a(\rho)$ should scale as $1/t_{col}$, where t_{col} is the mean time between collisions, set by the mean-free path (see **Appendix C**). So a should scale as ρ or $\rho^{\frac{1}{2}}$, which would be very different from our current assumption: the transition from the accumulation to the relaxation phase would be far smoother and especially the sharp peak in V would be very unlikely to occur. Similarly, all the V curves with different N would separate more progressively, leading to a shape closer to that of the above bimodal ABP model, so the improvement of the accumulation model would be partly lost.

A saturation in $a(\rho)$ could be easily introduced by assuming that the signal production is passivated for a time t_{min} upon activation: then $a(\rho)$ would saturate above ρ^* so that $t_{col}(\rho^*) = t_{min}$, introducing again a critical density ρ^* defining the time at which $V(t)$ leaves the initial carrying curve. However, again, the too smooth decrease of $a(\rho < \rho^*)$ would lead to a spread maximum in $V(t)$ and not a sharp peak (Fig. 4.43b).



(a) *(b)*

Figure 4.43: Accumulation model – smooth version.
(a) $R_g = \frac{2}{3}R$ *and (b)* V *as a function of time obtained from the accumulation model.* $a(\rho \geqslant \rho^*) = \beta V_m$, $a(\rho > \rho^*) = \frac{\beta V_m}{\rho^*}\rho$, *with* $\beta^{-1} = 80$ min, $V_m = 2.5$ µm.min^{-1}, $\rho^* = 9 \times 4$ cm^{-2}.

To sum up, the activity (persistence or speed) accumulation hypothesis seems promising, but it would need further development to assess its real accuracy. Moreover, it assumes a quite sophisticated intracellular process, which requires more experimental observation to the strength of its biological relevance.

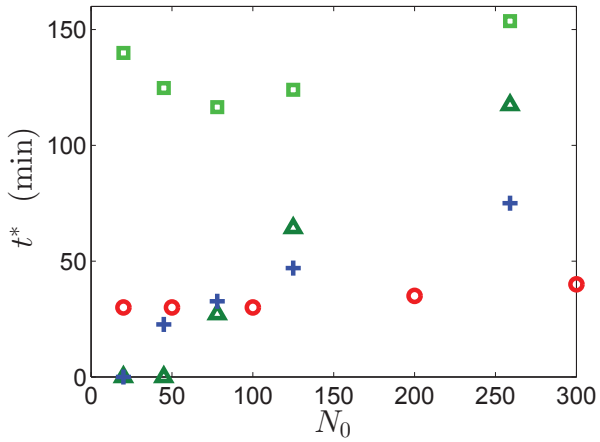


Figure 4.44: Summary: timing of the peak in radial velocity.
t^* as a function of N_0 from the experimental data (blue plus signs), the bimodal ABP model (red circles), the sharp accumulation model (dark green triangles) and the smooth accumuluation model (light green squares).

6 Conclusions and perspectives

In this chapter, we described an experimental observation of the spreading of cell colonies. In order to rationalise the analysis, the colonies were prepared using microfabrication techniques that allowed a precise control of the initial condition (colony shape, radius and density).

We noted that the spreading undergoes several regimes: at late times, it is dominated by proliferation – and slowing down due to a QSF effect – which meets the prediction of the FKPP equation; on a shorter time scale $t \leqslant 500$ min, however, a novel effect of contact enhancement of locomotion speeds up the process by a large amount. Thus, the cell-cell contacts act to increase the cell persistence, which in turn induces a local polarisation of the motion and a faster spreading of the population.

Various modelling approaches gave us a better insight into the dynamics of this effect. In particular, a model of bimodal, contact-activated active Brownian particles accounts well for this density dependent polarisation and speed up of the spreading. The main discrepancy resides in the scaling of the peak timing t^* with N_0, which could possibly be better described by an accumulation model (Fig. 4.44).

At the scale of the individual cell, this effect of CEL could be a particular manifestation of the CIL (for contact *inhibition* of locomotion), an effect that is well described in mammal cells [24]. In CIL, the protrusions in the contact zone are inhibited, hence leading to a new polarisation of the cell away from the contact. One could easily imagine that membrane

ruffling in the contact zone is inhibited for a while, inducing a transient more persistent motion (described in the bimodal ABP model of **Section 4.4**). The intracellular signals of polarisation and of motility activation could also be involved. For instance, the production of motility activating molecules could be induced transiently upon contact, leading to a contact-frequency-dependent accumulation of motility, accounted for in **Section 5.2**.

It seems fundamental to better understand these cell scale processes in order to go further in the interpretation of such results. To that end, specially designed experiments are needed. In particular, quantifying the effect of one or a series of contacts on single trajectories would allow to complete the picture. The main limitation in such experiments is that in two dimensions, when the cells are dilute enough to avoid multiple collisions, the contacts are actually too rare to yield enough statistics. Special experiments should thus be designed to force collisions, either by directing cells in a microsystem of by directly contacting a moving cell with another micro-manipulated one. Additional information on the molecular mechanisms underlying this interaction process could be brought by intracellular fluorescence imaging and molecular biology approaches.

At a larger scale, whatever its detailed mechanism, our observations demonstrated that CEL confers an increased invasive capacity to cells even in a relatively loose population. This could be at stake in leukocytes or highly metastatic cells that exhibit a motile behaviour close to *Dictyostelium*'s [118, 119]. Our experimental set-up and analysis framework could therefore be applied to these and other cell types to look for a similar CEL effect, or even to study other kinds of cell-cell interactions, for instance by adding cell-cell adhesion.

---

5. Dynamic aggregation at high cell density

Although *Dictyostelium* cells are well known for their aggregation upon sudden nutrient starvation, the behaviour of cells at high density in rich culture medium has raised relatively low interest until now. In this chapter, we report on experiments that aimed at understanding this limit of the density diagram. We observed an impressive phenomenon of dynamic aggregation: before confluence, the cells become more active and progressively gather into three-dimensional packed clusters, which are motile and exchange cells. We give a first description of the aggregation dynamics and quantify the aggregates' statistics, hence paving the way for the understanding of the exotic behaviours exhibited by *Dictyostelium* cells in extreme situations.

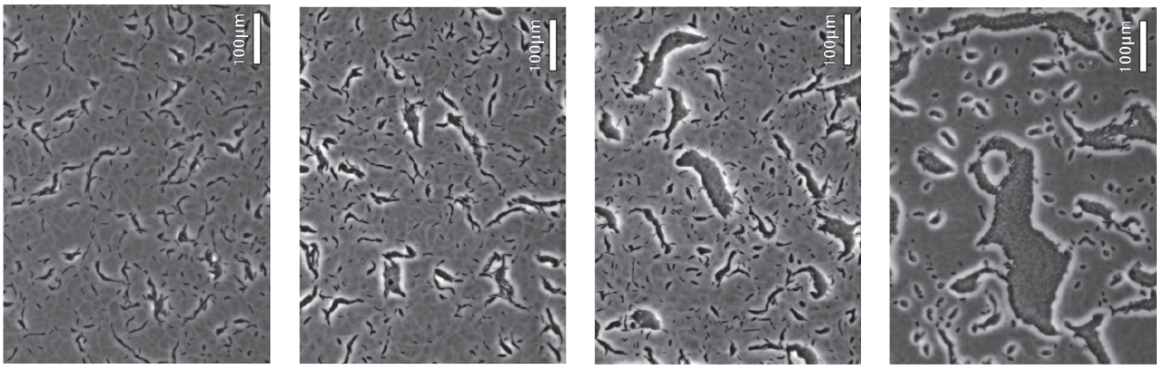
Contents

1 Introduction

We have seen in **Chapter 3** that density matters a lot. In particular, in the community of *Dictyostelium* biologists, common wisdom has it that one should refrain from letting the cells in culture reach too high densities, in order to avoid exotic or even pathological behaviours. As a consequence, we really wanted to let cells in culture reach high densities. Our observations about exotic, or even pathological behaviours are reported below. Yet, let us first discuss a little about the reasonable expectations of either biologists or physicists facing such a high density system.

For inert matter physicists, increasing the density usually means decreasing the activity. It typically involves a glassy transition, in which the dynamics slows down for kinetic reasons. Similar properties have been described in living tissues, and especially epithelia [27, 42] and studied theoretically for active matter [120]. However, some authors claim that in this particular case, the density should not be the driving parameter as the packing fraction is always 1, and they propose a model that highlights the role of various cell-cell adhesive interactions [121, 41].

Yet, below the jamming, increasing the density of active systems may also lead to richer phenomena. In the Vicsek model, alignment interactions leads to a transition to order with micro-phase separation of dense ordered bands [48, 50]. Such clustering associated to alignment has been described and modelled in rod-shaped bacteria [29, 30].



(a) *(b)* *(c)* *(d)*

Figure 5.1: Aggregation in Myxococcus xanthus (from [30]).
(a–d) Snapshots of clustering bacteria at various packing fractions: 0.06 *(a),* 0.1 *(b),* 0.16 *(c) and* 0.24 *(d). M. xanthus cells are rod-shaped, with length* $L = 6$ µm *and width* $W = 0.7$ µm*, and they align upon collision.*

Micro-phase separation also occurs without any adhesive or aligning interactions, by motility-induced phase separation (MIPS), when the

motility is decreased in denser zones. This happens for instance in homogeneous dense collections of hard-disks which average speed decreases because of steric hindrance. However, to our knowledge it has not yet been demonstrated to drive aggregation in any living system, even if it is a trendy hypothesis [122, 123].

2 Experimental observations

2.1 An overview of the phenomenon

The cells – DH1 strain – are seeded at 2×10^5 cm^{-2} in small home-made wells on a glass slide. Then they are simply imaged for up to one week. At first, the cells start proliferating – although not exponentially – and exhibit a very low motility, as expected at this high density (see **Chapter 3**). Then, they become suddenly more polarised and motile, leading to important local density fluctuations which progressively build up into dense, much likely three dimensional clusters (Fig. 5.2). These clusters grow in a very dynamic way: they include cells from the surrounding "dense gas", but also fuse with other clusters, melt sometimes, and move at random.

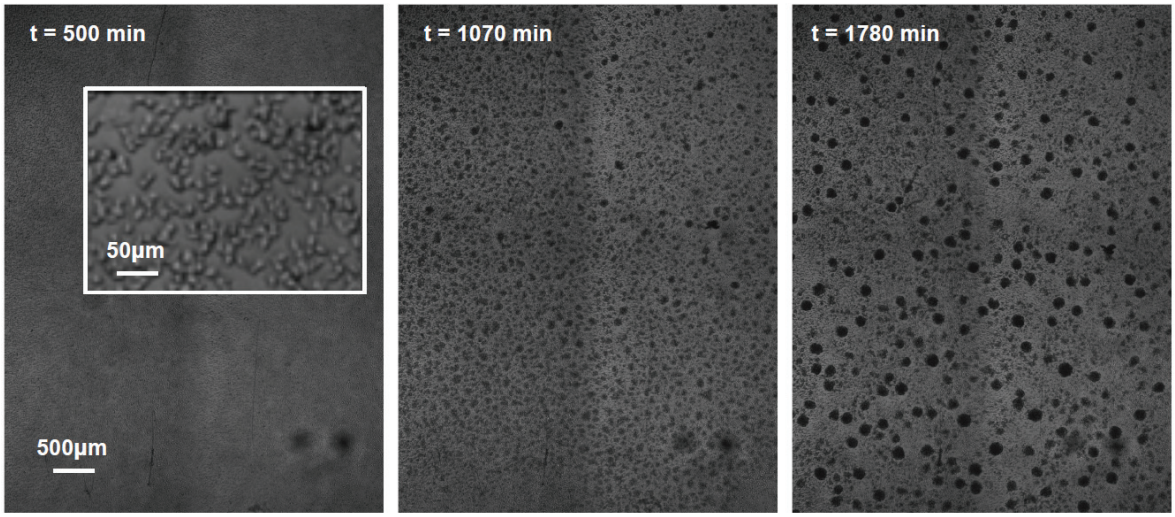


Figure 5.2: Cluster formation.
Starting from a homogeneous cell population at high density (left), the cells gather in small clusters (center) which grow with time to reach a stationary state (right). Left, inset: zoom in on the left picture. The cell density is high but below close packing.

We took advantage of the contrast between the dark aggregates and

the light gray background to segment the pictures (see Fig. 5.5c for an example of binarised image), and analyse the properties of the cell clusters.

The aggregation is characterised by a phase of rapid growth that lasts only a few hours, so that $\zeta(t)$, the fraction of the surface occupied by the aggregates, exhibits an almost sigmoid shape, with a sharp transition that separates the waiting phase from an almost stationary "aggregated" phase (Fig. 5.3). At the same time, the proliferation almost stops.

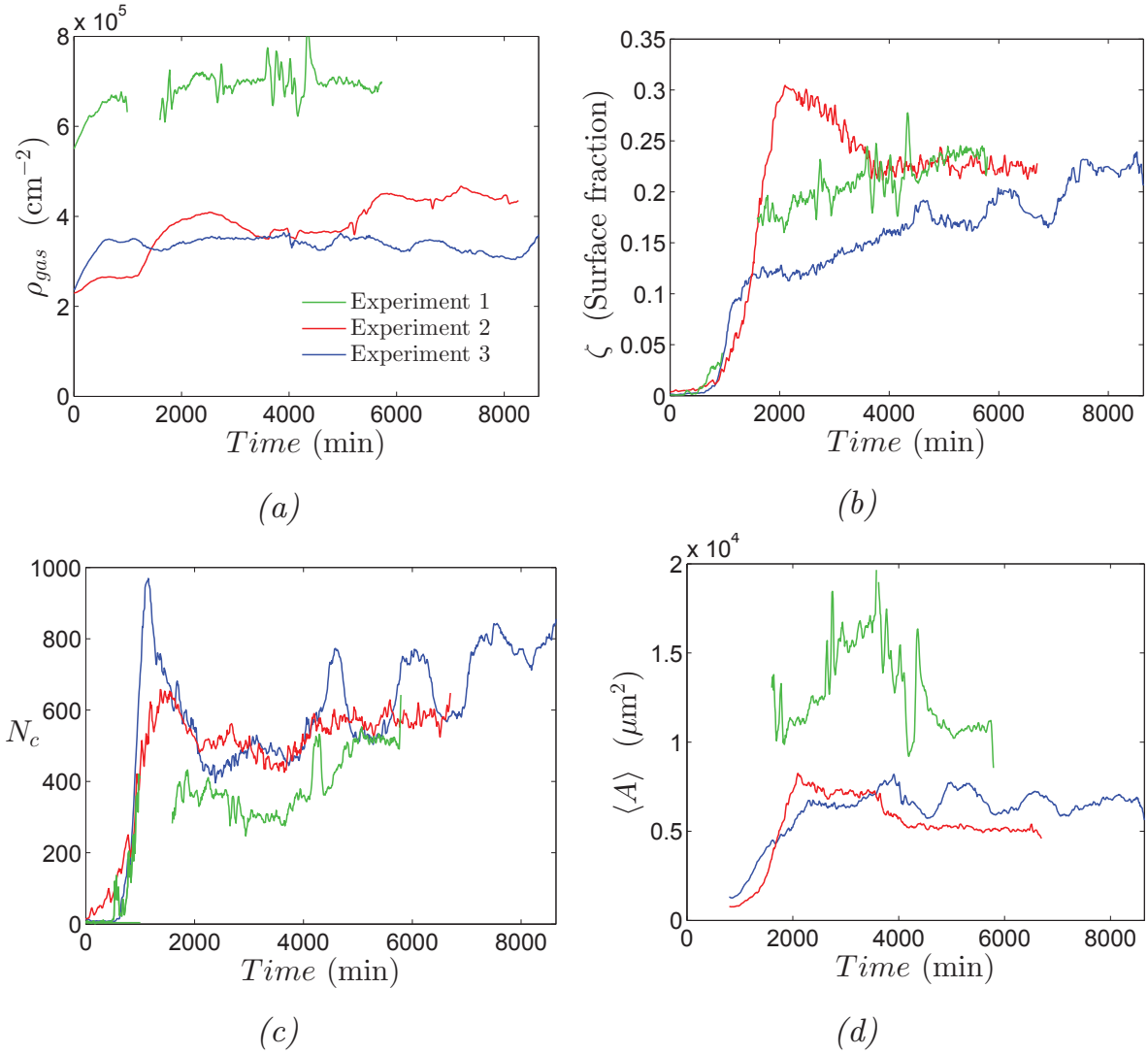


Figure 5.3: Aggregation kinetics.
(a) Cell density ρ_{gas} detected outside the aggregates, (b) surface fraction ζ occupied by the aggregates, (c) aggregate number N_c and (d) aggregate average area $\langle A \rangle$ as a function of time. Three independent experiments were aligned using the $N_c(t)$ curves.

A surrounding 'gas' of single cells. Around the aggregates, single cells keep behaving as in a homogeneous culture: they still move

seemingly at random on the surface, constituting what we term here a surrounding 'gas', in contrast with the condensed clusters. Interestingly, the cell density ρ_{gas} in the 'gas' remains constant, and relatively far from the close packing (Fig. 5.3a). The measurements seem to show a good agreement between experiments 2 and 3 (red and blue respectively) but not with experiment 1 (green). Yet one should not focus too much on the absolute values but rather on the time evolution: the aspect of pictures does not allow to really detect every single cell once, but the possible bias is expected to be constant over a whole experiment.

Kinetics of aggregation. To look deeper into the kinetics of aggregation, we computed both the cluster number N_c and average area $\langle A \rangle$ (Fig. 5.3c-d). We found that both increase from 0 to an almost stationary value, just as $\zeta(t)$ (Fig. 5.3). Yet, they exhibit slight differences. $N_c(t)$ has a very sharp initial increase and overtakes its aggregated phase value; then it relaxes slowly to the latter. By contrast, the onset of the growth of $\langle A \rangle(t)$ is delayed; then it increases more slowly, and it reaches its stationary value later than ζ and N_c. This shows that there is a first nucleation phase during which small clusters form; then these cluster grow by absorbing isolated cells from the surrounding "gas" and by fusing together, which explains the eventual decrease of N_c.

Pre-aggregation, dissolution, re-aggregation. In experiments run on shorter times and smaller fields of view, we observed an intriguing phenomenon: in some cases the cells first formed small transient clusters that melted completely before a new more extensive aggregation phase – likely the one described above – could start several hours later. The reason of this disintegration remains unclear. This might be related to the "breathing" phenomenon mentioned in **Section 2.2** below.

Spatial organisation. Whether the aggregation is globally or locally triggered is not obvious. In some experiments, it seemed to nucleate homogeneously at the scale of the field of view – several millimetres, still – but in some others we could observe an advancing front of aggregation. By computing N_c locally in adjacent strips, we could measure a local passage time and found a fluctuating front moving at almost constant speed (Fig. 5.4).

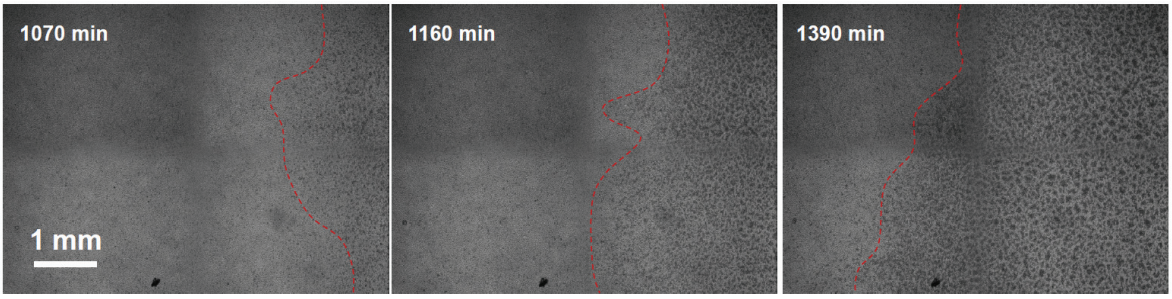


Figure 5.4: Aggregation front.
Small clusters first form on the right side of the field of view, then an aggregation front (the red dashed line is a guide to the eye) progresses to the left at a constant speed $c_{front} = 8.9$ µm.min^{-1}.

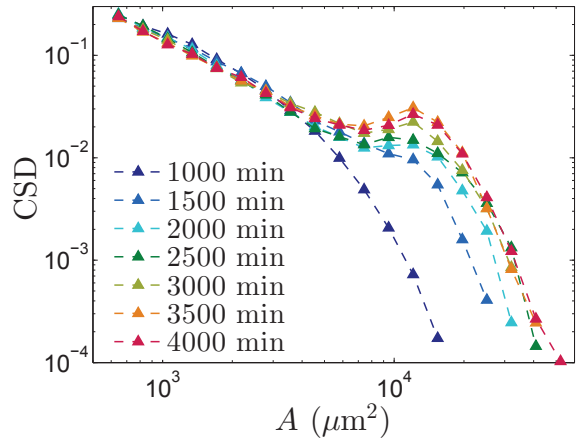
2.2 Aggregate size: statistics and dynamics

Cluster size distribution. We measured the cluster size distribution (CSD), defined as the relative frequency of aggregates of area A (Fig. 5.5). It evolves from an approximately exponential distribution at the onset of aggregation towards a specific stationary shape (Fig. 5.5a): a first power-law-like decrease is followed by a local maximum and a faster decay. The peak area defines the typical size $A^* = 1 - 1.5 \times 10^4$ µm^2 of numerous large and long-lived aggregates. The initial decay means that there is a large number of clusters of all sizes between 0 and A^*.

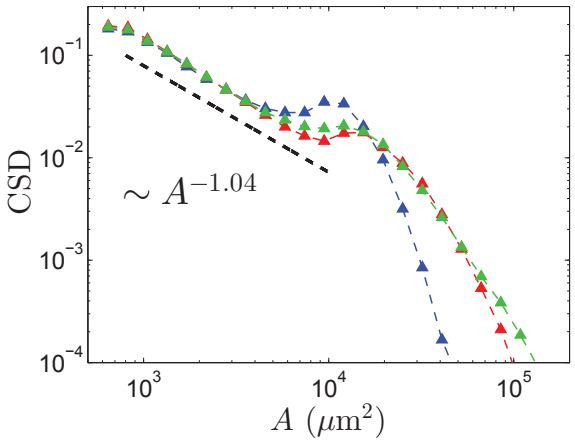
Interestingly, this distribution is well reproduced in three independent experiments made over more than a month (Fig. 5.5b), which shows that it is a robust property of this system.

Aggregate breathing. Even in the 'stationary' phase, the aggregates remain highly dynamic. In particular, they constantly seem to exchange cells, either with the surrounding 'gas' or with other aggregates. At a larger scale, the whole population of clusters breathes. This phenomenon is particularly marked in one of the experiments: we observe clear antiphase oscillations of the number of aggregates N_c and their average area $\langle A \rangle$ (Fig. 5.3, experiment 3, blue lines). In the CSD, it results in a depopulation of the large-scale peak at $A^* > 10^4$ µm in favour of more numerous clusters of intermediate size (Fig. 5.6).

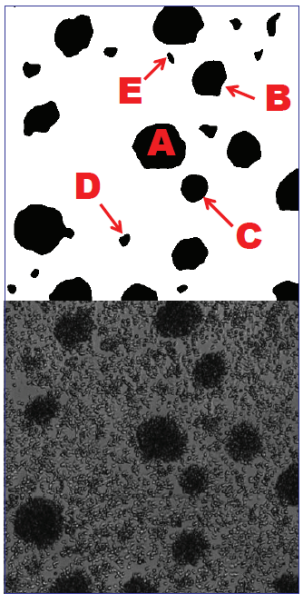
Interestingly, the CSD exhibits two distinct characteristic shapes corresponding to breathing in – fewer but larger aggregates, red up-pointing triangles – and out – numerous intermediate aggregates, blue down-pointing triangles.



(a)



(b)



(c)

Figure 5.5: Cluster size distributions.
(a) Evolution of the CSD with time in experiment 3 (in blue in (b) and in Fig. 5.3). (b) Average CSD for 2000 min $< t \leqslant$ 4000 min *for three independent experiments.*
(c) Binarisation process: the top image is the result of the unprocessed bottom picture. The lettered aggregates are examples of the variety of sizes: [A] 2×10^4 µm^2, [B] 1×10^4 µm^2, [C] 7×10^3 µm^2, [D] 1×10^3 µm^2, [D] 7×10^2 µm^2

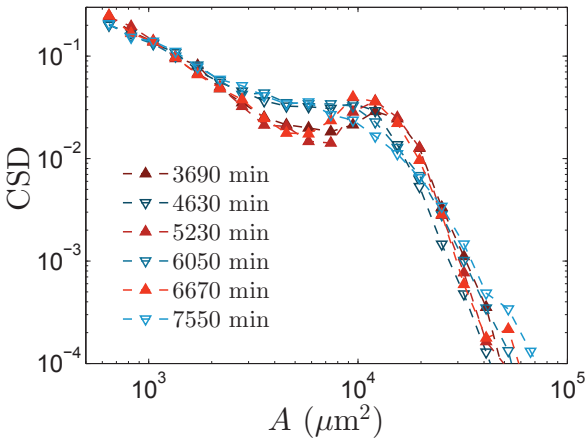


Figure 5.6: Aggregate breathing. CSD in experiment number 3, averaged over 100 min *around the hollows (red, solid up triangles) and peaks (blue, empty down triangles) in $N_c(t)$.*

3 Perspectives

The phenomenon that we have described here raises many questions, from both perspectives of the physiology of *Dictyostelium* and of the

physics of aggregating active systems.

Aggregation signal. Firstly, it is not clear whether the aggregation is an emergent property of the system or whether a signal is actually transmitted over the population to trigger a change of cellular behaviours. In the former case, an intrinsic parameter – such as the cell density or any characteristic of the motion – would reach a critical value leading to destabilisation of the homogeneous population; in the latter, the communication could occur by contact or *via* a change in the medium, *e.g.* the secretion or depletion of a molecule, or medium acidification. Of course, a mixture of the two is conceivable. For instance, a relative depletion of any nutrient could lead to a pre-starvation behaviour, involving an increase in motility, and the system's stability be sensitive to the cells' speed or persistence. We started experiments with a perfusion of fresh medium to test the hypothesis of medium-change-induction of the aggregation, but at that time the fluidic system we used was not reliable enough to produce convincing results.

It is even not clear whether the process nucleates locally or on a large scale. In two experiments (1 and 3, respectively blue and green lines in Fig. 5.3), there is a front of nucleation of the small clusters, while in the third experiment (red lines in Fig. 5.3) this nucleation step occurs simultaneously over the whole field of view, but then the formation of dark large aggregates starts on one side and propagates. In both situations the front comes from outside the field of view and extends further than the edges when it enters the image. This propagation could be the signature of a signal relay but it does not inform about the nature of this signal.

Finally, the constant density in the 'gas' is reminiscent of a condensed phase / gas coexistence, but ρ_{gas} is difficult to handle experimentally because of the proliferation, and at least some features – as the breathing phenomenon – seem to have a different origin.

Clustering mechanisms. In such a collection of moving agents, various mechanisms can lead to aggregation. The simplest one is the action of an attractive cell-cell interaction, which can take the form of chemoattraction but also of cell-cell adhesion. The interplay of the latter and the forces exerted on the substrate can stabilise more or less the aggregates. In principle, vegetative *Dictyostelium* cells do not express cell-cell adhesion proteins, but the long culture time at high density could trigger a prestarvation behaviour involving the expression of Contact Site

A, an important cell-cell adhesion protein expressed in the course of development [124] or other similar proteins. Experiments with knock-out cells, or with cells whose expression of prestarvation genes is tractable by fluorescence could give valuable information on this point.

Alternatively, active particles can cluster without explicit attraction: a motion that slows down in dense zones can lead to motility-induced phase separation (MIPS) [125]. The characteristics of the large aggregates, which are tightly packed, seemingly three-dimensional and motile, suggest that the cells inside adhere to each other, but MIPS could be at play in the early stage of aggregation.

Indeed, the aggregation involves two distinct morphologies of clusters: the smaller are two-dimensional and highly dynamic, especially at the onset of aggregation where they are similar to transient local density fluctuations; the larger look like three-dimensional dense structures, more comparable to spheroids. To better understand the phenomenon, it would be interesting to get a more precise picture of the aggregates' structure and of the single cell dynamics, in the clusters and in the 'gas'. The former could be achieved by using confocal imaging to get three-dimensional stacks of the aggregates, and at least higher magnification images. To that end, the main obstacle is that the appearance of the clusters depends on the magnification, and thus mapping higher magnification picture to our low magnification ones is not as straightforward as it could seem. Concerning the single cell dynamics, the experiment 1 was done with a mixture of GFP and RFP expressing cells (kindly provided by C. Nizak, ESPCI, Paris) in 98:2 ratio so that the RFP cells can be tracked individually. The aim was to separate the cells in the clusters from those in the 'gas', compare their motility and measure the entrance/exit rates in the aggregates (Fig. 5.7). However, for lack of additional experiments, these data were left on the side.

To sum up, various mechanisms, possibly triggered by various signals and control parameters, can lead to this rich dynamic aggregation behaviour. The main difficulty resides in the low control on the aggregation that is currently accessible. In that view, the spreading assay presented in **Chapter 4** could be a promising way, as the same clustering phenomenon was observed in the late high density stages of these experiments. Yet, although very preliminary, these results show that a broad class of rich exotic behaviours remains to be explored in biological systems near the onset of collectivity.

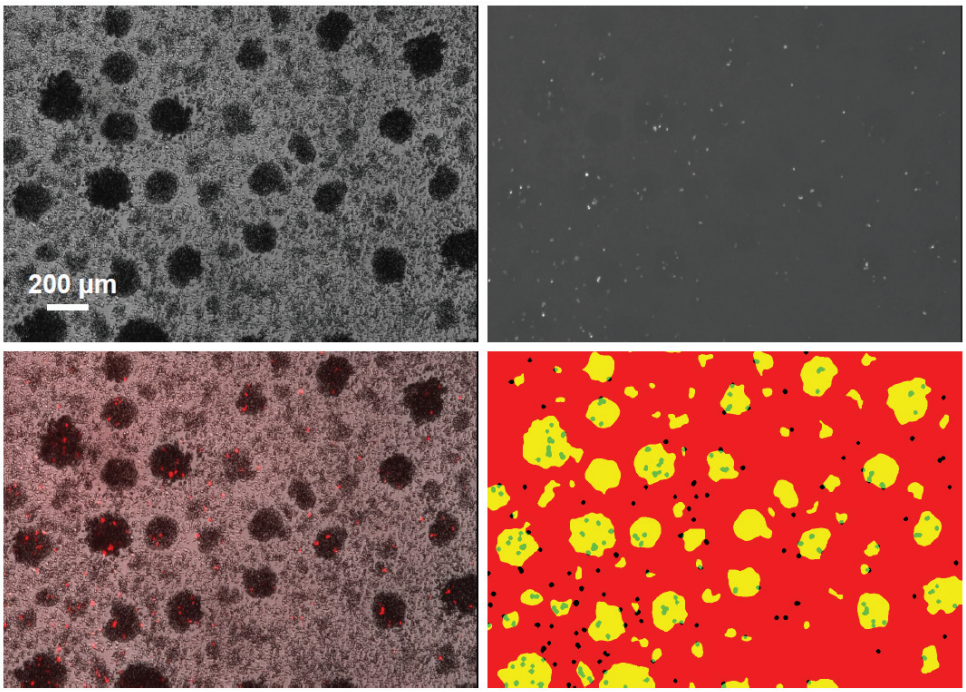


Figure 5.7: Tracking individual cells in the aggregates.
In experiment 1, GFP cells were mixed with RFP cells so that RFP cells are isolated. Bright-field (top-left), RFP channel (top-right), merge (bottom-left) and merged binarised images (bottom-right).

6. General conclusion

This thesis was based on three different experimental studies, reported in **Chapters 3, 4** and **5**, complemented with modelling when it could shed light on some principles driving the system. These experiments aimed at understanding how cells could regulate their motility as a group, by means of individual cell-cell interactions. To that end, we committed to the approach of physics, namely by choosing a simple microscopic observable and trying to relate its properties to the evolution of macroscopic variables. This involved a coarse graining procedure, whose details could in turn evidence the mechanisms linking the two scales.

To be more specific, we made the decision to work essentially on cell trajectories, hence considering the cells as particles whose internal dynamics was summed up in this apparently simple object. To be able to quantify the extent to which they were affected by the interactions, it was important to have a good idea of their properties in the absence of external cue. The review and preliminary measurements presented in **Chapter 2** showed that, even though this task itself was far from being completed, the cell trajectories could well be seen as an integrated manifestation of their internal activity. In particular, we stressed out the importance of the dynamics of the cell polarity axis, characterised by one or several persistence times.

In order to come and go between the individual and collective scales, active matter constitutes a very suitable framework. Indeed, it allows a quite straightforward integration of the experimental measurements into simple models of self-propelled particles, that describe closely the motion of polar individuals. Quantifying the effects of various classes of interactions is then made possible with retained generality. In **Chapter 4**, it allowed us to demonstrate the necessity of a Contact Enhancement of Locomotion (CEL) effect to account for the speed-up of early colony spreading, whereas classically accepted cell-cell interactions failed to it.

More generally, active matter has the potential to deal with interactions of all the kinds and functional forms that we have encountered during the course of this doctoral research. We pointed out three of them:

chemical sensing at a distance is the central point of **Chapter 3** and is much likely at stake in the triggering of the clustering in **Chapter 5**; a local 'repulsive' interaction, in the form of CEL, is responsible for the density sensitivity of colony spreading reported in **Chapter 4**; lastly, adhesive contacts probably play a driving role in the dynamic aggregation of **Chapter 5**.

It is remarkable that a single cell type is able to exhibit such various properties depending on the experimental context. To that extent, *Dictyostelium discoideum* has proven once again that it is a rich and powerful tool to uncover and decipher elementary biological functions, and especially regarding cell motility and emergent collective properties. The other side of the coin is that, should they be restricted to social amœbas, the results obtained with this organism would be of limited impact as such.

Therefore we expect this thesis to lay the foundation of research in more applied fields of cell biology. This is commonplace for findings initially made in *Dictyostelium* to be extended to leukocytes or even highly metastatic cells, that both present amœboid characteristics [126, 23]. Similarly, the interaction mechanisms that we have described here can reasonably be thought to matter in various situations. To our knowledge, quorum-sensing has not been reported so far to regulate the motility of other cell types, mainly because it has not been looked upon. However, there is no reason for it not to exist elsewhere. It can show itself to be particularly important *in vitro* as, when overlooked, it can bias the results in an unexplained fashion – this is actually the way it was first hypothesised. Besides, *in vivo*, cell populations that are scattered but still need to be able to reach any point in an extended zone could take advantage of such a regulation of the motility to achieve a good coverage. This includes *e.g.* immune cells or fibroblasts.

The effect of CEL is interesting for cells that may be more packed but where individuals can move freely over intermediate distances. Indeed it gives such a group the ability to spread more efficiently when its density is increased even in the absence of continual communication. It could be at play in the invasiveness of highly metastatic, loosely connected, tumour cells, or in microbial dispersion. Again it could also help immune cells to scatter from a possible prior meeting point.

Last, the dynamic breathing aggregation phenomenon undergone at high density contains certainly makes room for exciting discoveries on large-scale synchronisation and biphasic-equilibria-like phenomena that could occur in numerous physiological processes.

To conclude, this work is not an account of completed stand-alone research. Quite the opposite, I hope it can be considered as a working base for further findings in the broad and fascinating field of the emergent collective behaviours in living matter.

Appendices

A Bimodal rotational diffusion

Here we present results obtained by F. Detcheverry (Institut Lumière Matière, Université Claude Bernard Lyon 1, Villeurbanne) on the correlation function and the MSD of a particle undergoing bimodal rotational diffusion.

Assume a particle moves in the plane with velocity of constant magnitude v_0. Its orientation changes according to rotational diffusion, but with a coefficient that alternates between two values d_1 and d_2. The times spent in mode 1 and 2 are both exponentially distributed, with mean λ_1^{-1} and λ_2^{-1} respectively [1]. What are the properties of such a random motion?

The main quantity to determine is the angle correlation function

$$C_\theta(t'-t) = \langle \cos\left[\theta(t') - \theta(t)\right] \rangle, \tag{1}$$

which is equivalent to the velocity direction autocorrelation function (VDACF, see **Chapter 2**)

$$C(t'-t) = \langle \frac{\mathbf{v}(t') \cdot \mathbf{v}(t)}{\|\mathbf{v}(t')\| \|\mathbf{v}(t)\|} \rangle, \tag{2}$$

and its Laplace transform $C_\theta(s)$, from which one obtains the mean square displacement $m(t)$ and the diffusion coefficient from

$$m(s) = v_0^2 \, \frac{C_\theta(s)}{s^2}, \tag{3}$$

$$D = \frac{v_0^2}{4} \lim_{s \to 0} C_\theta(s). \tag{4}$$

One way to obtain $C_\theta(s)$ is to solve the problem

$$\begin{aligned} \partial_t p_1 &= d_1 \, \partial^2_{\theta\theta} p_1 - \lambda_1 p_1 + \lambda_2 p_2, \qquad & p_1(\theta, 0) = \phi_1 \delta(\theta), \\ \partial_t p_2 &= d_2 \, \partial^2_{\theta\theta} p_2 + \lambda_1 p_1 - \lambda_2 p_2, \qquad & p_2(\theta, 0) = (1 - \phi_1) \delta(\theta). \end{aligned} \tag{5}$$

Here $p_{i=1,2}(\theta, t)$ is the probability to be in mode i at time t, d_i is the rotational diffusion coefficient in mode i, and $\phi_1 = \lambda_2/(\lambda_1 + \lambda_2)$ is the fraction of time spent in mode 1. Introducing $p = p_1 + p_2$, $C_\theta(t)$ is given by

$$C_\theta(t) = \int_{-\pi}^{\pi} d\theta \, p(\theta, t) \cos\theta \tag{6}$$

[1] The value of d is thus a Telegraph process.

Using Laplace and Fourier transforms, one finds

$$C_\theta(s) = \frac{(\lambda_1+\lambda_2)^2 + \lambda_1(s+d_1) + \lambda_2(s+d_2)}{(\lambda_1+\lambda_2)\left[(s+d_1)(s+d_2) + \lambda_2(s+d_1) + \lambda_1(s+d_2)\right]}$$

In time domain one obtains for $C_\theta(t)$ the sum of two exponentials:

$$C_\theta(t) = Ce^{-\gamma t} + C'e^{-\gamma' t}, \tag{7}$$

with the notations

$$\begin{aligned}
\kappa^2 &= (d_1+d_2+\lambda_1+\lambda_2)^2 - 4(d_1d_2 + d_2\lambda_1 + d_1\lambda_2), && (8)\\
\kappa'^2 &= (d_1-d_2+\lambda_1-\lambda_2)^2 + 4\lambda_1\lambda_2, && (9)\\
\gamma &= (\kappa + d_1 + d_2 + \lambda_1 + \lambda_2)/2, && (10)\\
\gamma' &= \gamma - \kappa, && (11)\\
C &= -\frac{(d_1-d_2)(\lambda_1-\lambda_2) + (\lambda_1+\lambda_2)(\lambda_1+\lambda_2-\kappa')}{2\kappa(\lambda_1+\lambda_2)}, && (12)\\
C' &= \frac{(d_1-d_2)(\lambda_1-\lambda_2) + (\lambda_1+\lambda_2)(\lambda_1+\lambda_2+\kappa')}{2\kappa(\lambda_1+\lambda_2)}. && (13)
\end{aligned}$$

As a check, in the cases $d_1 = d_2$, $\lambda_2 \to \infty$ or $\lambda_1 \to 0$, one recovers the unimodal case and $C_\theta(t) = e^{-d_1 t}$.

Besides, the following inequalities hold:

$$C, C' \geqslant 0, \qquad d_1 \leqslant \gamma' \leqslant d_2 < \gamma, \tag{14}$$

showing that the slowest relaxation is intermediate between d_1 and d_2, and that both exponential are always decaying: $C_\theta(t)$ is thus a strictly decaying function and can not have a minimum. Finally, the diffusion coefficient D is

$$\frac{4D}{v_0^2} = \frac{\lambda_1 d_1 + \lambda_2 d_2 + (\lambda_1+\lambda_2)^2}{(\lambda_1+\lambda_2)(d_1d_2 + \lambda_2 d_1 + \lambda_1 d_2)} \tag{15}$$

In the limit $d_2 \to \infty$, *i.e.* when all directional persistence is lost in mode 2, all expressions greatly simplify

$$\begin{aligned}
\gamma &= d_2, \quad \gamma' = d_1 + \lambda_1, && (16)\\
C &= \frac{\lambda_1}{\lambda_1+\lambda_2}, \quad C' = \frac{\lambda_2}{\lambda_1+\lambda_2}, && (17)\\
\frac{4D}{v_0^2} &= \frac{\lambda_2}{(\lambda_1+\lambda_2)(d_1+\lambda_1)}, && (18)
\end{aligned}$$

where only the first term in the expansion has been retained. In this case, $C_\theta(t)$ exhibit a rapid drop over a time $\sim d_2^{-1}$, followed by a slowest decay whose constant $\gamma' = d_1 + \lambda_1$ is independent of d_2. This is the limit used in **Appendix B** to calibrate a bimodal analysis procedure.

The model used in **Chapter 2, Section 4.2** and in **Chapter 4, Section 4.3** to model the effect of CEL is another special case of bimodal rotational diffusion with $d_2^{-1} = 0$.

As final remark, note that the motion is different from what would be observed with two populations of cells, each remaining in a given mode. In that case, the quantities are obtained by an average over the two populations, for instance

$$C_\theta(t) = \overline{e^{-d_i t}}, \quad \frac{4D}{v_0^2} = \overline{d_i^{-1}}, \quad \overline{X_i} = \phi_1 X_1 + (1-\phi_1) X_2, \tag{19}$$

which are distinct from the expressions above.

B Towards bimodal analysis

Assuming that a given trajectory is bimodal, one might want to recover the underlying parameters – d_1, d_2, the rotational diffusion constants and λ_1, λ_2, the mode transition rates, in the case of a bimodal ABP. This is possible in theory using the fit of the autocorrelation function. However, inverting the expressions for τ_1, τ_2, ϕ_1 and ϕ_2 is far from straightforward, and the relative imprecision of the measurements could be misleading to precisely estimate the four parameters at the same time. This is why we prefer to go no further than the estimation of ϕ_1 and ϕ_2 with fixed τ_1 and τ_2, which at least gives a reliable picture of the relative amounts of short and long relaxation times.

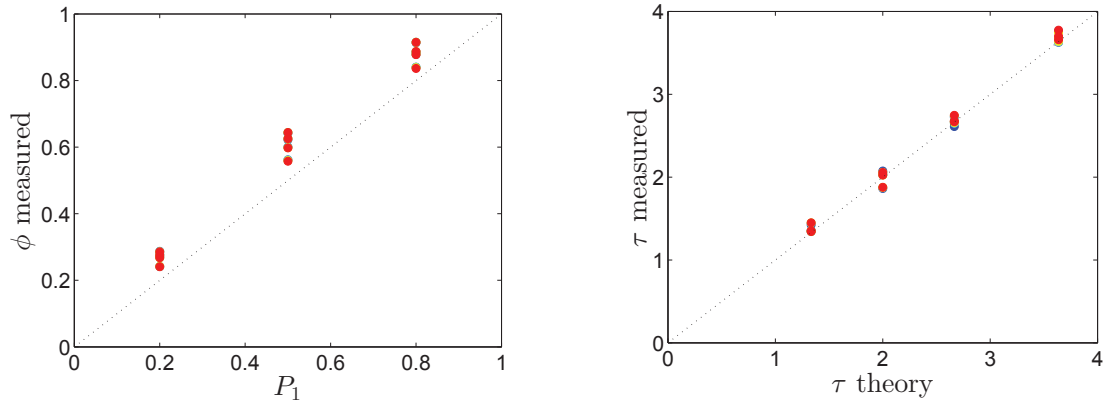
To go further, it could be interesting to be able to segment the trajectories into steps of each mode. This would allow to precisely study the statistics of each mode. However, until now, the only proposed analysis schemes have relied on arbitrary criteria. The typical way is to define one or a few observables that may characterise the current mode: consecutive turning angles [71, 18], or local MSD exponent together with the local turning angle variance [74]. In case the two modes are really different, setting the suitable threshold might not be too difficult and the modes can be detected quite reliably. Yet the situation in which bimodal analysis is more valuable is when the mode detection is more subtle. Of course, even in this case, this kind of procedure always yields a result. The question is, how informative is this result? And also, what is the influence of an arbitrarily set criterion for mode allocation?

Bimodal active-passive Brownian particles The Coefficient of Movement Efficiency that we defined in **Section 3.2** seems to be a good candidate for mode detection. Indeed, it is close to one when the motion is persistent, and to 0 when it is random. To evaluate a bimodal analysis scheme based on the CME, we simulated particles alternating between an ABP mode '1' with persistence $\tau_p = 4$ min and speed $v_0 = 5$ µm.min^{-1} and a diffusion mode '0' with diffusion constant D_t (no persistence at all). Both modes are exponentially distributed with mean step durations τ_1 and τ_0 respectively. The data were discretised so that $\delta t = 20$ s and $\delta x = 0.67$ µm. The control parameter used is the proportion of time spent in mode 1, $P_1 = \frac{\tau_1}{\tau_1+\tau_0}$, which arises from the velocity autocorrelation function

$$C(t) = \phi e^{-\frac{t}{\tau}} \tag{20}$$

where

$$\begin{cases} \phi = P_1 \\ \tau = \frac{\tau_p \tau_1}{\tau_p + \tau_1} \end{cases} \tag{21}$$



(a) *(b)*

Figure B1: Parameters of the velocity autocorrelation function. (a) Measured ϕ and (b) τ as a function of their theoretical predictions, for various τ_1 and τ_0. Measurements from fits of the correlation function over 5, 10, 15 and 20 minutes (blue, light blue, green and red respectively) are completely superimposed.

We first checked whether we could recover well this parameters by fitting the autocorrelation function. The results presented in Figure B1 shows a good agreement with the predictions, the slight overestimation of ϕ being well understood in the perspective of **Section 4.1**.

Distribution of the CME. To characterise the distributions, we plotted the inverse cumulative distribution function $\Phi(c) = \int_c^1 \mathbb{P}(c')dc'$ that describes the proportion of data points with CME$> c$ (see **Chapter 2** for the definition of the CME). This function is noted Φ on purpose as we will look for c such that $\Phi(c) = \phi$ (the ϕ in the correlation function) later. This function decreases from 1 at $c = 0$ to 0 at $c = 1$. A fast decrease denotes a peak in the probability density function, while a flat zone stands for a hole. A "good" shape in our expectations would be a sigmoid decrease with a very flat part in the middle, whose value would be close to P_1. It would mean that there is a clear separation between the peak in the CME distribution due to the random part of the trajectories and the one due to the persistent part.

Of course, the shape also depends on the Δt that is chosen to compute the CME. One expects that if $\Delta_t << t_p$, where t_p is the characteristic decorrelation time, then both peaks will be located near 1. In the opposite limit, both will be next to 0 because $\lim_{\Delta t \to +\infty} CME_{\Delta t} = 0$ for both modes.

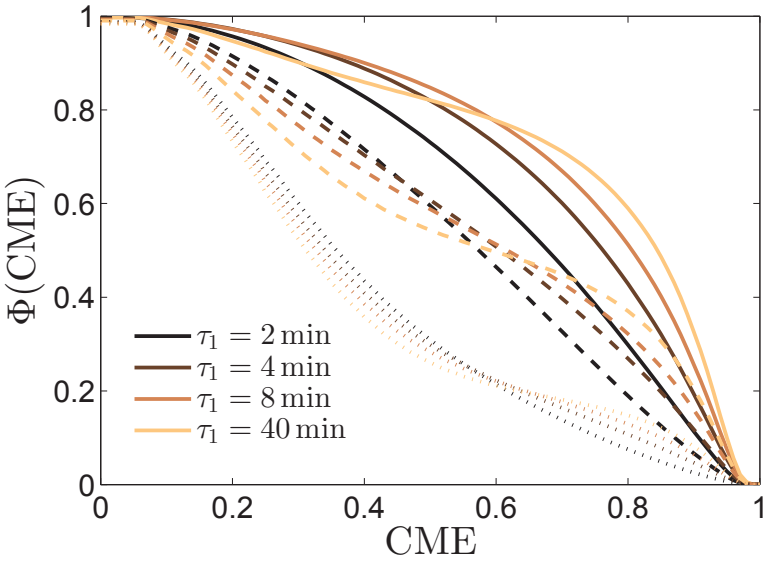


Figure B2: Cumulative distribution of CMEs measured at $\Delta t = 4$ min *for* $P_1 \in \{0.8, 0.5, 0.2\}$ *(full, dashed and pointed lines respectively) and* $\tau_1 \in \{2, 4, 8, 40\}$ min *(from darker to lighter brown).*

In Figure B2 we display the Φ cumulative function measured at $\Delta t = \tau_p = 4$ min with various input parameters. One notices three bundles of curves that correspond to the three values of P_1. The greater P_1 is, the later $\Phi(CME)$ decreases. In each bundle the shape evolves from single curvature to sigmoidal shape when τ_1 is increased. This is an effect of time scale separation. Indeed, when the steps in trajectories are short compared to, or of the same order as Δt, the CME values arise typically from mixtures of both modes, leading to a very smooth distribution. Conversely, when the steps in trajectories are longer, the time window typically corresponds to one or the other mode, hence two better separated distributions of CME.

One can see several question emerging from this single figure:

(i) How does the shape of Φ depend on P_1 ?

(ii) How does the shape depend on τ_1 ?

(iii) Does the shape depend on τ_p itself ?

(iv) How to optimize Δt ?

The answer of the question (i) looks quite clear: the greater P_1 is, the later the function decreases, giving rise to shapes that evolve from - globally - convex to - globally - concave.

The role of τ_1 appears also quite clearly, as explained above: when the time scales - τ_1 and Δt - are mixed, the CME distribution essentially consists in a wide bump, while when these scales are well separated, the distribution is bimodal, leading to a flat zone in $\Phi(c)$, hopefully so that this "almost-plateau" value is next to P_1 (this would mean that one can distinguish the modes according to their CME value easily).

The plots in figure B3 give a good idea of what is a "good separation of timescales". Indeed the curves tend to a limit shape which is almost reached at $\tau_1 = 10\tau_p$ – *ie* for steps in mode 1 10-fold longer than the persistence time – while at $\tau_1 = 2\tau_p$ the plateau is still not visible at all.

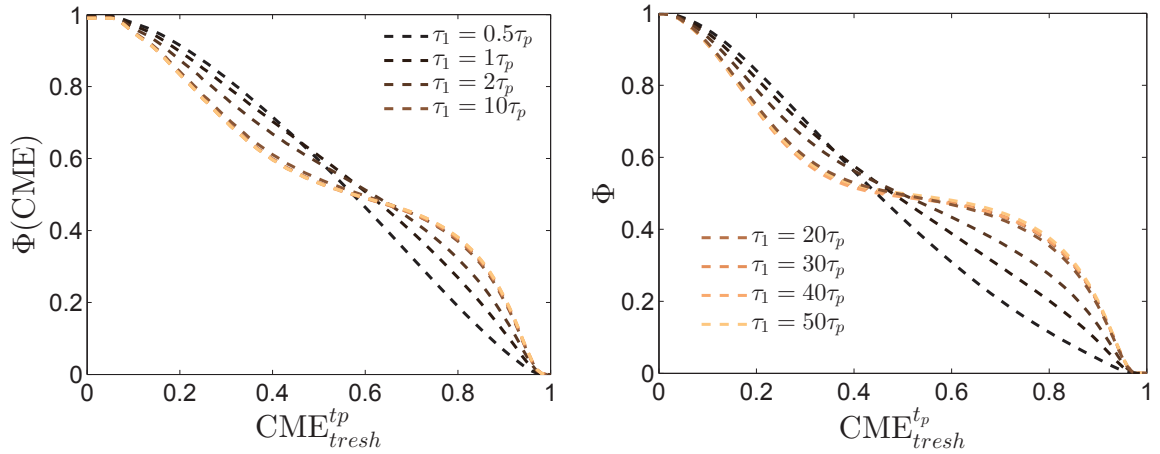


Figure B3: Cumulative distribution of CMEs measured at $\Delta t = t_p$ *for* $t_p = 4$ min *(left) and* $t_p = 8$ min *(right),* $P_1 = 0.5$ *and* $\tau_1 \in \{0.5, 1, 2, 10, 20, 30, 40, 50\} \times \tau_p$ *(from darker to lighter brown).*

The role of τ_p may be – a little – more subtle. Thus, when everything can be normalized by τ_p, different τ_p values should yield the same results as long as the other – normalized – parameters match. Yet, it is again not the case here, as can be seen on figure B3. This is mainly a discretisation effect, which again shows itself to be non-negligible at all. Indeed the pixel size defines a minimal correlation length – all the trajectories are at least straight over the pixel size – and the frame rate defines a new time scale – which can be seen, similarly, as a minimal correlation time. Thus, it is easier to isolate random portions of the trajectory when δt, this "correlation time" of mode 0 is well separated from τ_p. One can

check that binning the data at $\tau_p = 8$ min twice in space and time yields the same Φ(CME) as for $\tau_p = 4$ min (data not shown).

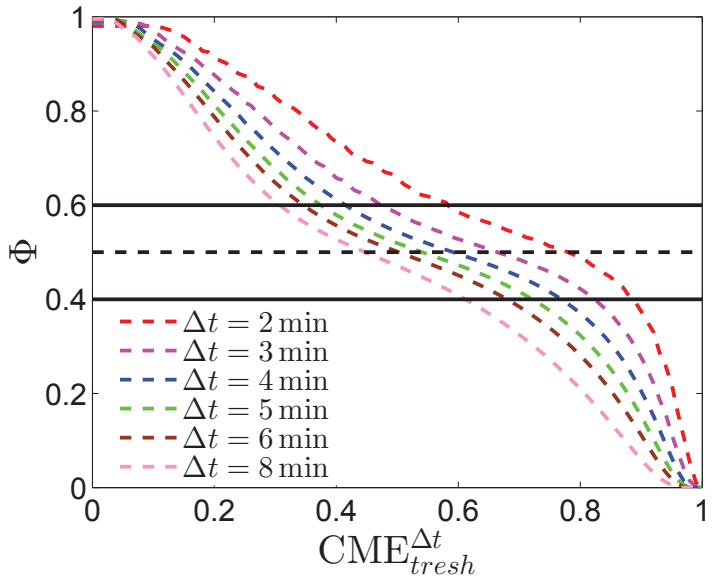


Figure B4: Cumulative distribution of the CME measured at $\Delta t \in \{2, 3, 4, 5, 6, 8\}$ min *(red, fuschia, blue, green, brown and pink respectively) for* $\tau_p = 4$ min, $P_1 = 0.5$ *and* $\tau_1 = 40$ min. *The horizontal lines are at* $\Phi = 0.5$ *(solid),* 0.4 *and* 0.6 *(dashed).*

The last question is the optimization of Δt, the time frame used to compute the CME. This is not a very easy one for several reasons. We have just seen that it might be interesting to have $\Delta t >> \delta t$. However Δt should be kept of the order of τ, otherwise the persistent portions will appear random.

Figure B4 shows that choosing $\Delta t \sim \tau$ is probably the right choice. Indeed, this is around this value that the plateau is the most marked, and also the values of Φ in this zone are then correctly located near P_1.

Last, one should keep in mind that we do not know τ_p *a priori* from the data, but rather only τ (see Equation 20). As a consequence, this value should be taken as reference to set Δt, even though the Φ curves for various τ_1 collapse better onto each other when one uses $\Delta_t = \tau_p$ than $\Delta t = \tau$.

Bimodal analysis scheme. The properties of the Φ function are now quite well described. From this we can conclude – not very satisfactorily – that a universal CME_{thresh} value that would separate the random and persistent portions of trajectories cannot be defined.

Instead, one should use a systematic procedure to set its value for every

dataset, knowing that the output will probably be of better quality when the Φ(CME) curve exhibits a flat part in the middle.

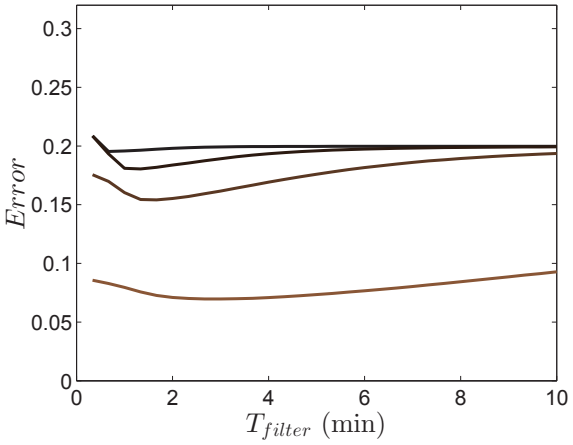
The procedure we propose is the following:

(i) From the raw trajectories, measure ϕ and τ *via* the velocity direction correlation function

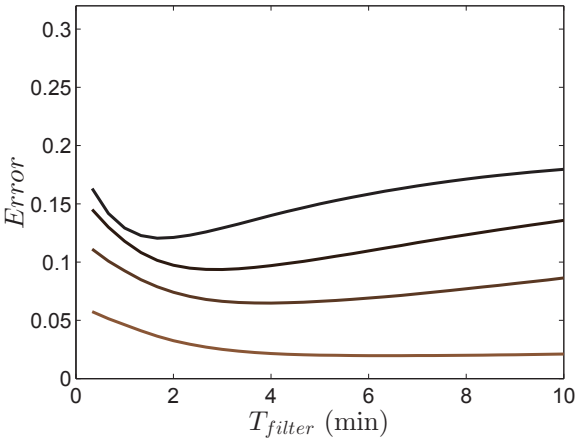
(ii) Use these values to compute CME_τ and find CME* so that $\Phi(\text{CME}^*_\tau) = \phi$.

(iii) Allocate the data points a mode (1 or 0) according to their CME$_\tau$, using CME* as threshold.

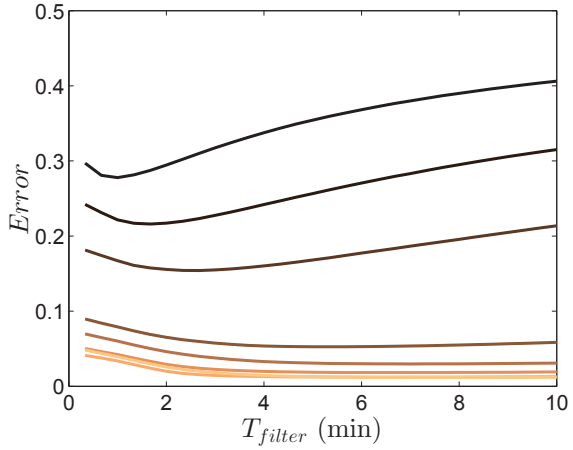
(iv) (Optionnal step) Filter the obtained binary data so that no portion is shorter than T_{filter}.



(a)



(b)



(c)

Figure B5: Error rate as a function of the minimum step duration T_{filter}. Mode allocation using the systematic procedure described above for (a) $P_1 = 0.8$, (b) $P_1 = 0.2$, with $\tau_1 \in \{2, 4, 8, 10\}$ min and (c) $P_1 = 0.5$ with $\tau_1 \in \{2, 4, 8, 10, 20, 30, 40, 50\}$ min. τ_1 is color-coded, from dark to light.

Allocation error. The principle of such an analysis is to allocate a $Mode$, either 0 or 1 to each time point in the trajectory, aiming at reproducing reliably original sequence ($Mode_{real}$). We estimated the quality of this process by computing the error defined as:

$$Error = \langle |Mode_{real} - Mode_{Analysis}| \rangle \tag{22}$$

The results of this estimation are plotted in figure B5. To better see the impact of this procedure, the y-axes where limited to the "random allocation error" defined by $E_{rand} = 2P_1(1-P_1)$, which is the error that would result from randomly allocating the mode to the data points.

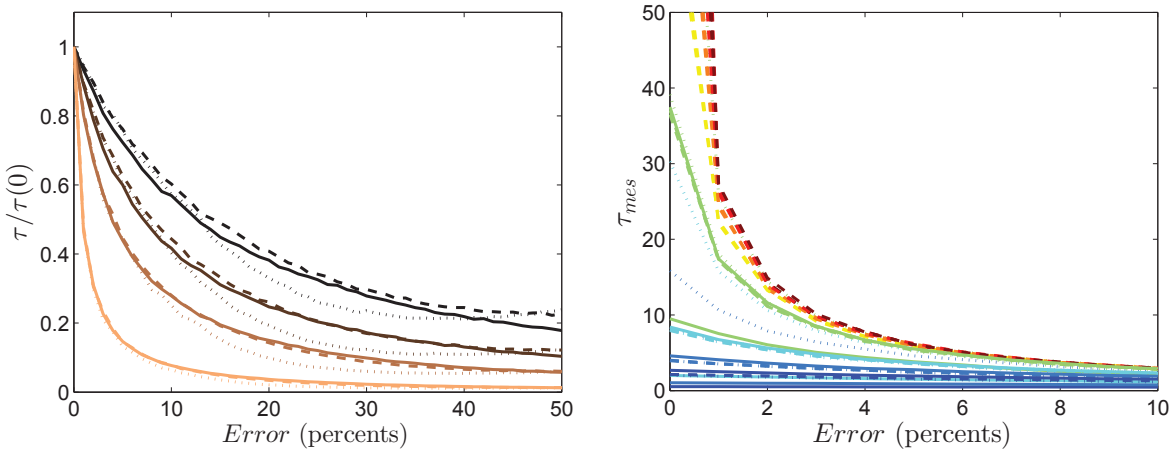
Again, for $\tau_1 \sim \tau_p$ the results do not look exceptionally good compared to the random allocation. Yet it does not mean that the measurement of the statistics – τ_1, τ_0, P_1 – is not greatly improved compared to random allocation!

Error in time constants. With random allocation, the probability of finding a persistent portion of duration δt is $\mathbb{P}_1 = 1 - \phi^2$, and for $n > 1$, the probability of finding a segment of duration $n\delta t$ is given by $\mathbb{P}_n = \frac{1}{Z} e^{-n(1-\phi)}$ where Z is a normalization factor.

As this is an exponential distribution, one should be very careful about the values of τ_1 and τ_0 that arise from this procedure, because the shape of the distribution can be misleading. Moreover, as illustrated in figure B6a, adding only little random errors to a sequence of zeros and ones with exponential distributions of the segment lengths affects the measured τ_1 and τ_0 a lot.

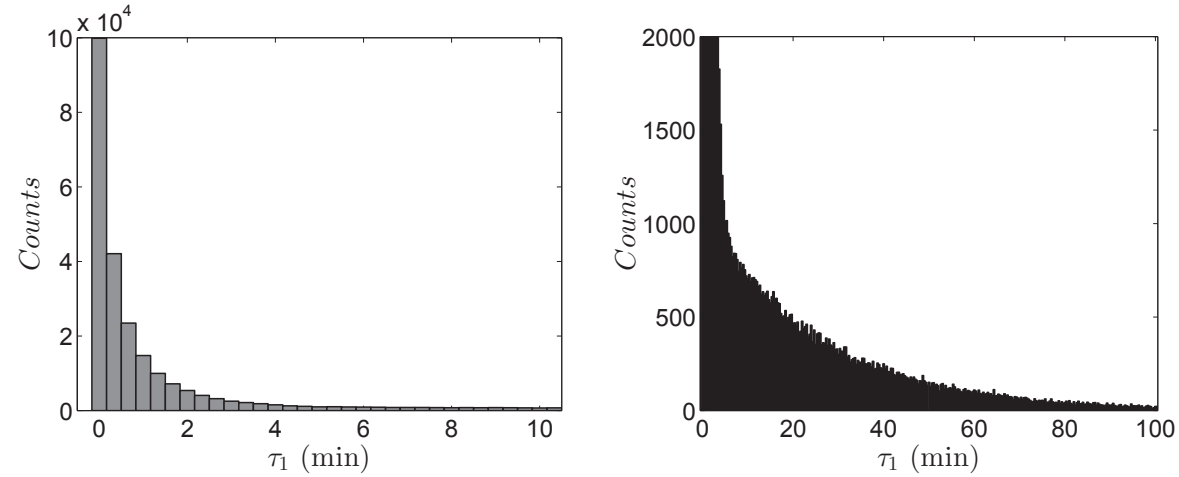
In particular, with a rate of added random errors of only 5%, all the measured outputs are below 10 min, whatever the actual value (Fig. B6b). It means that only a few random errors masks completely the actual underlying distribution of step duration by shifting the distribution to the left. It is because each error cuts a segment of length T into to smaller pieces: thus, adding only as many errors as there are steps of both modes will decrease the mean step duration by approximately a factor 2!

Finally, the τ_1 distribution that is obtained is a mixture of a rapid decay due to errors only and a slower one which results to the original segment cut by the errors (Fig. B7). It could be possible to extract information from this kind of distribution, but more work is needed in order to achieve precise and general analyses on this basis, and the finish line seems way further than the scope of this manuscript.



(a) (b)

Figure B6: Effect of random errors on τ_1 and τ_0.
The sequences of zeros and ones denoting the current mode were taken directly from the simulations' output. Then various amounts of random point errors was added to test the robustness of the step durations' statistics. (a) Normalized measured τ_1 for $\tau_1 \in \{2, 4, 8, 40\}$ min (from dark to light brown) and $P_1 \in \{0.2, 0.5, 0.8\}$ min (pointed, dashed and full lines respectively). (b) Measured τ_1 and τ_0 for all conditions with $\tau_p = 4$ min. No matter their true value, a few percents of error bring them down easily under 10 min



(a) (b)

Figure B7: Distribution of persistent step lengths after bimodal analysis. (a) Distribution of step lengths τ_1. (b) Same distribution: with a zoom in of the y-axis and a zoom out of the x-axis, a second larger time scale emerges.

Continuously evolving persistence. An alternative way to deal with non-constant persistence in single trajectories is a continuously

evolving underlying variable, rather than a two-states switching quantity. Metzner *et al.* have recently developed a method based on Bayesian inference to estimate the "instantaneous" persistence of a trajectory.

To sum up, they assume an underlying order 1 auto-regressive process (AR1-process) that is the discrete equivalent of an Ornstein-Uhlenbeck process:

$$\mathbf{v}_{t+1} = q_t \mathbf{v}_t + a_t \boldsymbol{\eta} \tag{23}$$

where $\mathbf{v}_t$ is the velocity vector at time t, $\boldsymbol{\eta}$ is a Gaussian white noise with unit variance, q_t is a persistence parameter and a_t an activity parameter. Their method allows them to recover q_t and a_t. In principle it could be applied directly to bimodal ABPs, which would yield a bimodal distribution of q_t and a_t. The technique could also be adapted to the ABP model by working on the direction of motion θ rather than on $\mathbf{v}$, resulting in a $D_r(t)$ function for all trajectories. In this scope D_r should be considered as distributed variable rather than as a constant parameter. Although promising, however, this approach is still very preliminary in the context of our data.

C A mean-field description of Contact Enhancement of Locomotion

In this part, we start analysing a mean-field model of CEL, based on the coarse-grained equations of the simulations in **Chapter 4, Section 4.4**.

Let us recall the classical ABP model on which the two-dimensional simulations were built:

$$\begin{cases} \dot{r}_i = v_0.\mathbf{e}_{\theta_\mathbf{i}} \\ \dot{\theta}_i = \sqrt{2D_r}\tilde{\eta}_i(t) \end{cases} \tag{24}$$

where $\tilde{\eta}_i(t)$ is a Gaussian white noise with zero mean and unit variance.

It has been shown [52] that the mean-field behaviour of the density ρ and the polarisation $\mathbf{p}$ obeys the following set of equations:

$$\begin{cases} \partial_t \rho = -\boldsymbol{\nabla}.(v_0\mathbf{p}) \\ \partial_t \mathbf{p} = -D_r\mathbf{p} - \frac{1}{2}\boldsymbol{\nabla}(v_0\rho) \end{cases} \tag{25}$$

In a homogeneous density field, the excluded volume interactions modify the equations only slightly:

$$\begin{cases} \partial_t \rho = -\boldsymbol{\nabla}.(v(\rho)\mathbf{p}) \\ \partial_t \mathbf{p} = -D_r\mathbf{p} - \frac{1}{2}\boldsymbol{\nabla}(v(\rho)\rho) \end{cases} \tag{26}$$

where $v(\rho) = v_0.(1 - f\frac{\rho}{\rho_m})$ is the effective local speed, which is decreased because particle motion is impeded during collisions ($f \approx 0.9$, see **Chapter 1**).

We used this mean-field system as a starting point for the continuous description of our system.

Should the hard-core repulsion be taken into account? The first thing to do is to estimate the effect of the $v(\rho)$ dependency. We solved the system (26) numerically with various N_0. It shows that the spreading rate decreases with the density (Fig.C1), contrary to the experimental

case. The importance of this effect could be predicted as for $\rho = 4.8 \times 10^5$ cm^{-2}, $v = 0.7v_0$, hence $D = 0.7^2 D_0 = 0.5D_0$, where D is the effective diffusion constant and D_0 the diffusion constant in the absence of interactions.

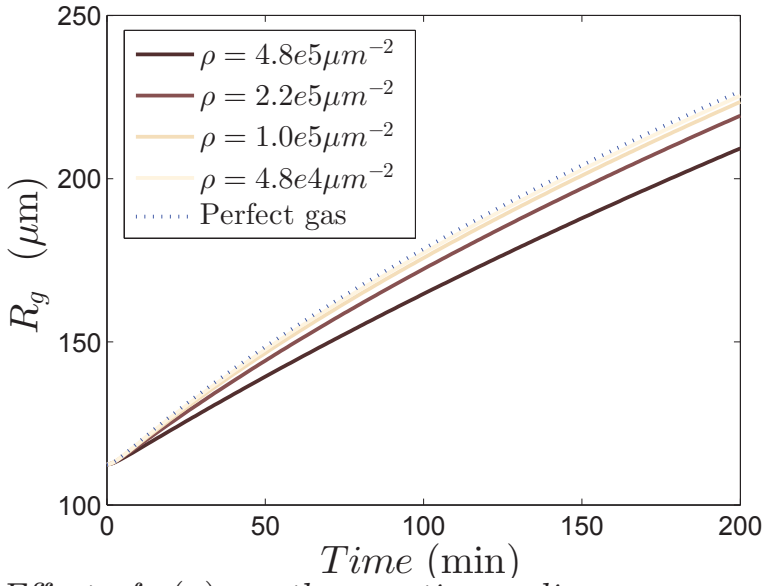


Figure C1: Effect of $v(\rho)$ on the gyration radius.
$R_g(t)$ for various cell numbers with the $v(\rho)$ dependency (solid lines) and limit with no interaction (dashed line). $\rho_m = 1.4 \times 10^6$ cm^{-2}, $v_0 =$ 7 µm.min^{-1}.

However, the particle-based simulations show no effect of the density when only hard-core repulsion is considered. It is probably because the aforementioned $v(\rho)$ is valid only for a uniform density field. In our situations with high gradients, other terms could play a role. For instance, integrating naively both this $v(\rho)$ and the effect described by Bruna and Chapman [116] yields

$$\frac{D}{D_0} \sim (1 + 4\frac{\rho}{\rho_m})(1 - f\frac{\rho}{\rho_m})^2, \tag{27}$$

or accounting for this by an effective speed,

$$\frac{v}{v_0} \sim (1 - f\frac{\rho}{\rho_m})\sqrt{1 + 4\frac{\rho}{\rho_m}}. \tag{28}$$

For $\rho/\rho_m \leqslant 0.5$, both these expressions take values close to 1. Thus, keeping a constant v_0 seems to be the best choice in our situation, which is also good news in terms of computing efficiency.

Peak of polarisation and persistence. We solved numerically the system (25) with an initial condition consistent with the experiments, $v_0 \in \{3; 5; 7\}$ µm.min^{-1} and $\tau_p = 2^n$ with $n \in [-1\ 6]$. All the obtained $v_r(t)$ exhibit a peak, but there is not trivial scaling between the peak timing t^* and the persistence time $\tau_p = D_r^{-1}$. Conversely, we could notice an interesting relation between the normalised peak timing t^*/τ_p and the normalised peak radial velocity v_r^*/v_0 (Fig. C2). Its significance remains unclear, but this is a new proof that the experimental observations cannot be accounted for by a constant speed and a constant persistence time, as both v_r^* and t^* increase with $< N_0 >$.

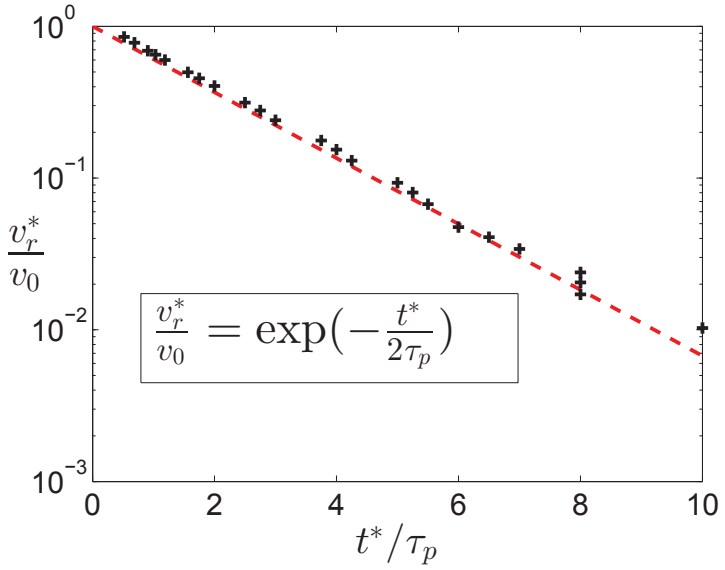


Figure C2: The normalised peak radial velocity v_r^*/v_0 scales exponentially with the normalised peak timing t^*/τ_p (black plus signs: simulations, dashed red line: exponential fit).

The basic two-state model. Now we consider ABPs that switch from their "natural" mode 1 upon collision to an "excited" or "escape" mode 2. The first mode is characterized by persistence time τ_1, while the second one is characterized by a higher persistence time, τ_2. This mode relaxes to mode 1 with a rate ω. The collision rate depends on the densities and is denoted t_{col}^{-1}. This gives rise to a system of four coupled equations.

$$\begin{cases} \partial_t \rho_1 = -\boldsymbol{\nabla}.(v_1 \mathbf{p}_1) - \frac{\rho_1}{t_{col}} + \omega \rho_2 \\ \partial_t \mathbf{p}_1 = -\frac{\mathbf{p}_1}{\tau_1} - \frac{1}{2}\boldsymbol{\nabla}(v_1 \rho_1) - \frac{\mathbf{p}_1}{t_{col}} + \omega \mathbf{p}_2 \\ \partial_t \rho_2 = -\boldsymbol{\nabla}.(v_2 \mathbf{p}_2) + \frac{\rho_1}{t_{col}} - \omega \rho_2 \\ \partial_t \mathbf{p}_2 = -\frac{\mathbf{p}_2}{\tau_2} - \frac{1}{2}\boldsymbol{\nabla}(v_2 \rho_2) + \frac{\mathbf{p}_1}{t_{col}} - \omega \mathbf{p}_2 \end{cases} \tag{29}$$

The speed of i^{th} mode is denoted v_i to let the possibility that the propul-

sion speed also depends on the motion mode.

Collision rate. In this model, the dynamics is controlled by the collision rate. *A priori* the exact value of the mean time between collisions is a complicated function of ρ_1, ρ_2, $\mathbf{p}_1$ and $\mathbf{p}_2$. To estimate this is itself a big question. However one can also try to find a correct approximation. Let us first consider simple ABPs that follow the equations 24 in the high and low density limits.

At high density the mean free path is shorter than the persistence length. Between two collisions, the particle motion is thus ballistic. Equating the mean free path and the distance covered between two collisions yields the collision time:

$$t_{col} = t_{bal} = \frac{1}{v\rho^{\frac{1}{2}}} \tag{30}$$

At low density, the motion is diffusive, hence

$$t_{col} = t_{dif} = \frac{1}{2v^2\tau\rho} \tag{31}$$

The curves for these two times cross at $\rho^* = \frac{1}{4v^2\tau^2}$, which defines the order of magnitude of the transition from one to the other regime (Fig. C3). In this monomodal model, a function that admits expressions (30) and (31) as high and low density limits and $t_{col}(\rho^*) \sim \tau$ should be a good approximation, for instance $t_{col} = t_{bal}.e^{-\frac{\rho}{\rho^*}} + t_{dif}.(1 - e^{-\frac{\rho}{\rho^*}})$ or $t_{col}^{-1} = t_{bal}^{-1}.e^{-\frac{\rho}{\rho^*}} + t_{dif}^{-1}.(1 - e^{-\frac{\rho}{\rho^*}})$.

In the experiments, taking as typical parameter values $v = 5$ µm.min^{-1} and $\tau = 4$ min yields $\rho^* = 6.25 \times 10^4$ cm^{-2}, so that at least at short times most of the experimental conditions are in the ballistic regime.

Of course, the bimodal model makes the estimation of t_{col} a little bit more complicated. Nevertheless, when $v_1 \sim v_2$ it should not change too much.

We solved this set of equations for various starting densities with $t_{col} = t_{dif}$, $v_1 = 3$ µm.min^{-1}, $\tau_1 = 2$ min, $v_2 = 7$ µm.min^{-1}, $\tau_2^{-1} = 0$ (ballistic mode) and $\omega^{-1} = 10$ min. The results exhibit the expected behaviour with a density-dependent spreading rate and peak of v_r (Fig. C4). Using $v_1 < v_2$ allows a better separation between the different N_0, but t^* is still roughly the same for all conditions.

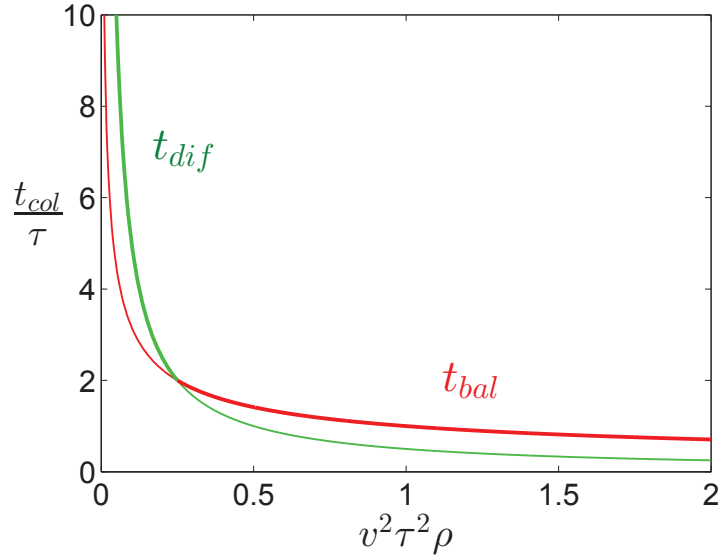
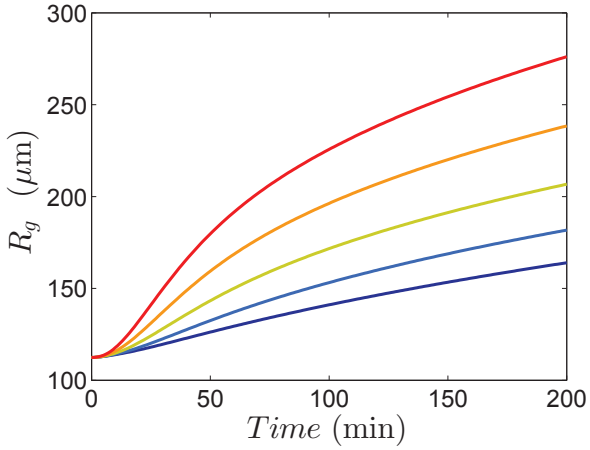
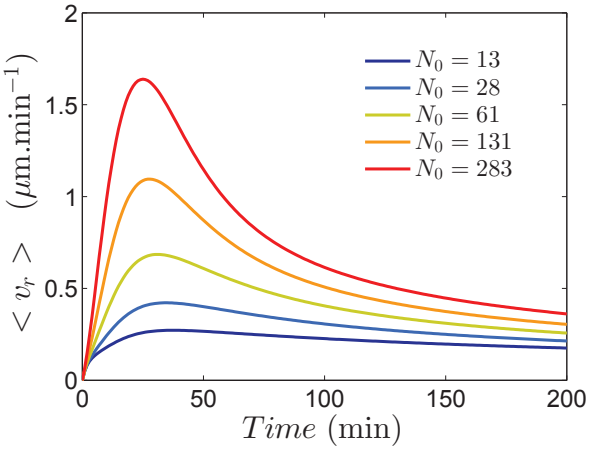


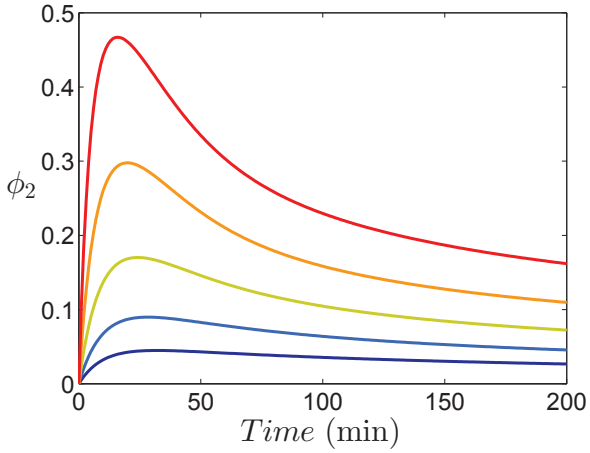
Figure C3: Diffusive (green) and ballistic limits of t_{col}. The lines are thicker in the regime where they should apply.



(a)



(b)



(c)

Figure C4: Results of the numerical solving of the mean-field CEL model.
(a)R_g as a function of time for various N_0. (b) $< v_r >$ as a function of time. (c) ϕ_2, measured as the proportion of cells in mode 2, as a function of time. It exhibits a density-dependent peak which is earlier than the $< v_r >$ peak.

References

[1] Tomita Vasilica Stirbat, Sham Tlili, Thibault Houver, Jean Paul Rieu, Catherine Barentin, and Hélène Delanoë-Ayari. Multicellular aggregates: a model system for tissue rheology. *The European physical journal. E, Soft matter*, 36(8):84, 2013.

[2] Anna Bagorda and Carole a Parent. Eukaryotic chemotaxis at a glance. *Journal of cell science*, 121(Pt 16):2621–2624, 2008.

[3] J. Fred Hess, Kenji Oosawa, Nachum Kaplan, and M. I. Simon. Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell*, 53(1):79–87, 1988.

[4] Jacques Monod. The growth of bacterial cultures. *Annual Reviews of Microbiology*, 3:371–394, 1949.

[5] Ewa Zlotek, Sylvain Monnier, Giovanni Capello, Maël Le Berre, and Matthieu Piel. Optical volume and mass measurements show that cells swell during mitosis. *The Journal of Cell Biology*, 211(4):765–74, 2015.

[6] Pierre-François Verhulst. Notice sur la loi que la population suit dans son accroissement. In *Correspondance mathématique et physique*, chapter 10, pages 113–121. 1838.

[7] S A McDonald and A J Durston. The cell cycle and sorting behaviour in Dictyostelium discoideum. *Journal of cell science*, 66:195–204, 1984.

[8] C J Weijer, G Duschl, and C N David. A revision of the Dictyostelium discoideum cell cycle. *Journal of cell science*, 70:111–131, 1984.

[9] Debra A Brock and Richard H Gomer. A secreted factor represses cell proliferation in Dictyostelium. *Development*, 132(20):4553–62, 2005.

[10] Deenadayalan Bakthavatsalam, Debra A Brock, N Neda Nikravan, Kevin D Houston, R Diane Hatton, and Richard H Gomer. The secreted Dictyostelium protein CfaD is a chalone. *Journal of cell science*, 121(Pt 15):2473–2480, 2008.

[11] Richard H. Gomer, Wonhee Jang, and Derrick Brazill. Cell density sensing and size determination. *Development Growth and Differentiation*, 53(4):482–494, 2011.

[12] Mélanie Chabaud, Mélina L. Heuzé, Marine Bretou, Pablo Vargas, Paolo Maiuri, Paola Solanes, Mathieu Maurin, Emmanuel Terriac, Maël Le Berre, Danielle Lankar, Tristan Piolot, Robert S. Adelstein, Yingfan Zhang, Michael Sixt, Jordan Jacobelli, Olivier Bénichou, Raphaël Voituriez, Matthieu Piel, and Ana-Maria Lennon-Duménil. Cell migration and antigen capture are antagonistic processes coupled by myosin II in dendritic cells. *Nature communications*, 6(May):7526, 2015.

[13] Pablo Vargas, Paolo Maiuri, Marine Bretou, Pablo J Sáez, Paolo Pierobon, Mathieu Maurin, Mélanie Chabaud, Danielle Lankar, Dorian Obino, Emmanuel Terriac, Matthew Raab, Hawa-racine Thiam, Thomas Brocker, Susan M Kitchen-Goosen, Arthur S Alberts, Praveen Sunareni, Sheng Xia, Rong Li, Raphael Voituriez, Matthieu Piel, and Ana-Maria Lennon-Duménil. Innate control of actin nucleation determines two distinct migration behaviours in dendritic cells. *Nature cell biology*, 18(1):43–53, 2016.

[14] Peter Friedl and Darren Gilmour. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology*, 10(7):445–457, 2009.

[15] Leonard Bosgraaf and Peter J M Van Haastert. The Ordered Extension of Pseudopodia by Amoeboid Cells in the Absence of External Cues. *PLoS ONE*, 4(4), 2009.

[16] Paolo Maiuri, Jean F. Rupprecht, Stefan Wieser, Verena Ruprecht, Olivier Bénichou, Nicolas Carpi, Mathieu Coppey, Simon De Beco, Nir Gov, Carl-Philipp Heisenberg, Carolina Lage Crespo, Franziska Lautenschlaeger, Maël Le Berre, Ana-Maria Lennon-Dumenil, Matthew Raab, Hawa-Racine Thiam, Matthieu Piel, Michael Sixt, and Raphaël Voituriez. Actin Flows Mediate a Universal Coupling between Cell Speed and Cell Persistence. *Cell*, 161(2):374–386, 2015.

[17] G.E. Uhlenbeck and L.S. Ornstein. On the theory of the Brownian motion. *Journal of the Physical Society of Japan*, 36(5):823–841, 1930.

[18] Laurent Golé, Charlotte Rivière, Yoshinori Hayakawa, and Jean Paul Rieu. A quorum-sensing factor in vegetative Dictyostelium Discoideum cells revealed by quantitative migration analysis. *PLoS ONE*, 6(11):1–9, 2011.

[19] M. T. Keating and J. T. Bonner. Negative chemotaxis in cellular slime molds. *Journal of Bacteriology*, 130(1):144–147, 1977.

[20] Jonathan Hickson, S. Diane Yamada, Jonathan Berger, John Alverdy, James O'Keefe, Bonnie Bassler, and Carrie Rinker-Schaeffer. Societal interactions in ovarian cancer metastasis: A quorum-sensing hypothesis. *Clinical and Experimental Metastasis*, 26(1):67–76, 2009.

[21] Evelyn F. Keller and Lee A. Segel. Model for chemotaxis. *Journal of Theoretical Biology*, 30(2):225–234, 1971.

[22] Jonathan E Phillips and Richard H Gomer. A secreted protein is an endogenous chemorepellant in Dictyostelium discoideum. *Proceedings of the National Academy of Sciences of the United States of America*, 109(27):10990–5, 2012.

[23] Sarah E Herlihy, Darrell Pilling, Anu S Maharjan, and Richard H Gomer. Dipeptidyl peptidase IV is a human and murine neutrophil chemorepellent. *Journal of immunology (Baltimore, Md. : 1950)*, 190(12):6468–77, 2013.

[24] Roberto Mayor and Carlos Carmona-Fontaine. Keeping in touch with contact inhibition of locomotion. *Trends in Cell Biology*, 20(6):319–328, 2010.

[25] Nicolas Borghi, Maria Sorokina, Olga G Shcherbakova, William I Weis, Beth L Pruitt, W James Nelson, and Alexander R Dunn. E-cadherin is under constitutive actomyosin-generated tension that is increased at cell-cell contacts upon externally applied stretch. *Proceedings of the National Academy of Sciences of the United States of America*, 109(31):12568–73, 2012.

[26] Dhananjay T Tambe, C Corey Hardin, Thomas E Angelini, Kavitha Rajendran, Chan Young Park, Xavier Serra-Picamal, Enhua H Zhou, Muhammad H Zaman, James P Butler, David A Weitz, Jeffrey J Fredberg, and Xavier Trepat. Collective cell

guidance by cooperative intercellular forces. *Nature materials*, 10(6):469–75, 2011.

[27] Thomas E. Angelini, Edouard Hannezo, Xavier Trepat, Jeffrey J. Fredberg, and David A. Weitz. Cell migration driven by cooperative substrate deformation patterns. *Physical Review Letters*, 104(16):1–4, 2010.

[28] M. Reffay, L. Petitjean, S. Coscoy, E. Grasland-Mongrain, F. Amblard, A. Buguin, and P. Silberzan. Orientation and polarity in collectively migrating cell structures: Statics and dynamics. *Biophysical Journal*, 100(11):2566–2575, 2011.

[29] H P Zhang, Avraham Be'er, E-L Florin, and Harry L Swinney. Collective motion and density fluctuations in bacterial colonies. *Proceedings of the National Academy of Sciences of the United States of America*, 107(31):13626–13630, 2010.

[30] Fernando Peruani, Jornn Starruss, Vladimir Jakovljevic, Lotte Sogaard-Andersen, Andreas Deutsch, and Markus Bär. Collective motion and nonequilibrium cluster formation in colonies of gliding bacteria. *Physical Review Letters*, 108(9):1–5, 2012.

[31] G Duclos, S Garcia, H G Yevick, and P Silberzan. Perfect nematic order in confined monolayers of spindle-shaped cells. *Soft Matter*, 10(14):2346–2353, 2014.

[32] M. Abercrombie and Joan E.M. Heaysman. Observations on the social behaviour of cells in tissue culture. *Experimental Cell Research*, 5(1):111–131, 1953.

[33] Carlos Carmona-Fontaine, Helen K Matthews, Sei Kuriyama, Mauricio Moreno, Graham A Dunn, Maddy Parsons, Claudio D Stern, and Roberto Mayor. Contact inhibition of locomotion in vivo controls neural crest directional migration. *Nature*, 456(7224):957–61, 2008.

[34] Elena Scarpa, Alice Roycroft, Eric Theveneau, Emmanuel Terriac, Matthieu Piel, and Roberto Mayor. A novel method to study contact inhibition of locomotion using micropatterned substrates. *Biology open*, 2(9):901–6, 2013.

[35] Ravi A Desai, Smitha B Gopal, Sophia Chen, and Christopher S. Chen. Contact inhibition of locomotion probabilities drive solitary versus collective cell migration. *Journal of the Royal Society, Interface / the Royal Society*, 10(88):20130717, 2013.

[36] Sri Ram Krishna Vedula, Man Chun Leong, Tan Lei Lai, Pascal Hersen, Alexandre J Kabla, Chwee Teck Lim, and Benoît Ladoux. Emerging modes of collective cell migration induced by geometrical constraints. *Proceedings of the National Academy of Sciences of the United States of America*, 109(32):12974–9, 2012.

[37] Camila Londono, M Jimena Loureiro, Benjamin Slater, Petra B Lücker, John Soleas, Suthamathy Sathananthan, J Stewart Aitchison, Alexandre J Kabla, and Alison P McGuigan. Nonautonomous contact guidance signaling during collective cell migration. *Proceedings of the National Academy of Sciences of the United States of America*, 111(5):1807–1812, 2014.

[38] Man Chun Leong, Sri Ram Krishna Vedula, Chwee Teck Lim, and Benoît Ladoux. Geometrical constraints and physical crowding direct collective migration of fibroblasts. *Communicative and Integrative Biology*, 6(2):37–41, 2013.

[39] Søren Vedel, Savaş Tay, Darius M Johnston, Henrik Bruus, and Stephen R Quake. Migration of cells in a social context. *Proceedings of the National Academy of Sciences of the United States of America*, 110(1):129–34, 2013.

[40] T E Angelini, E Hannezo, X Trepat, M Marquez, J J Fredberg, and D A Weitz. Glass-like dynamics of collective cell migration. *Proceedings of the National Academy of Sciences of the United States of America*, 108(12):4714–4719, 2011.

[41] Dapeng Bi, J. H. Lopez, J. M. Schwarz, and M. Lisa Manning. A density-independent rigidity transition in biological tissues. *Nature Physics*, 11(12):1074–1079, 2015.

[42] Simon Garcia, Edouard Hannezo, Jens Elgeti, Jean-François Joanny, Pascal Silberzan, and Nir S Gov. Physics of active jamming during collective cellular motion in a monolayer. *Proceedings of the National Academy of Sciences of the United States of America*, 112(50):15314–15319, 2015.

[43] Janina Lange, Vera Auernheimer, Pamela L. Strissel, and Wolfgang H. Goldmann. Influence of focal adhesion kinase on the mechanical behavior of cell populations. *Biochemical and Biophysical Research Communications*, 436(2):246–251, 2013.

[44] A.-K. Marel, A. Piera Alberola, and J. O. Rädler. Proliferation and Collective Migration of Small Cell Groups Released From Circular Patches. *Biophysical Reviews and Letters*, 07(01n02):15–28, 2012.

[45] Eric Theveneau, Lorena Marchant, Sei Kuriyama, Mazhar Gull, Barbara Moepps, Maddy Parsons, and Roberto Mayor. Collective Chemotaxis Requires Contact-Dependent Cell Polarity. *Developmental Cell*, 19(1):39–53, 2010.

[46] Mae L. Woods, Carlos Carmona-Fontaine, Chris P. Barnes, Iain D. Couzin, Roberto Mayor, and Karen M. Page. Directional collective cell migration emerges as a property of cell interactions. *PLoS ONE*, 9(9), 2014.

[47] Eric Theveneau, Benjamin Steventon, Elena Scarpa, Simon Garcia, Xavier Trepat, Andrea Streit, and Roberto Mayor. Chase-and-run between adjacent cell populations promotes directional collective migration. *Nature Cell Biology*, 15(7):763–772, 2013.

[48] T Vicsek, A Czirok, Eshel Ben-Jacob, I Cohen, and O Shochet. A novel type of phase transition in a system of self-driven particels. *Physical Review Letters*, 75(6), 1995.

[49] Guillaume Grégoire and Hugues Chaté. Onset of collective and cohesive motion. *Physical review letters*, 92(2):025702, 2004.

[50] Alexandre P. Solon, Hugues Chaté, and Julien Tailleur. From phase to microphase separation in flocking models: The essential role of nonequilibrium fluctuations. *Physical Review Letters*, 114(6):1–5, 2015.

[51] Jean Baptiste Caussin, Alexandre Solon, Anton Peshkov, Hugues Chaté, Thierry Dauxois, Julien Tailleur, Vincenzo Vitelli, and Denis Bartolo. Emergent spatial structures in flocking models: A dynamical system insight. *Physical Review Letters*, 112(14):1–5, 2014.

[52] Alexandre P. Solon, M. E. Cates, and Julien Tailleur. Active brownian particles and run-and-tumble particles: a comparative study. *European Physical Journal, Special Topics*, 224(7):1231–1262, 2015.

[53] Alexandre P Solon, Y Fily, A Baskaran, M E Cates, Y Kafri, M Kardar, and J Tailleur. Pressure is not a state function for generic active fluids. *Nature Physics*, 11(August):1–7, 2015.

[54] Alexandre P. Solon, Joakim Stenhammar, Raphael Wittkowski, Mehran Kardar, Yariv Kafri, Michael E. Cates, and Julien Tailleur. Pressure and Phase Equilibria in Interacting Active Brownian Spheres. *Physical Review Letters*, 114(19):1–6, 2015.

[55] Yaouen Fily and M. Cristina Marchetti. Athermal phase separation of self-propelled particles with no alignment. *Physical Review Letters*, 108(23):1–5, 2012.

[56] Christophe Deroulers, Marine Aubert, Mathilde Badoual, and Basil Grammaticos. Modeling tumor cell migration: From microscopic to macroscopic models. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 79(3):1–14, 2009.

[57] Gianluca Ascolani, Mathilde Badoual, and Christophe Deroulers. Exclusion processes: Short-range correlations induced by adhesion and contact interactions. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(1), 2013.

[58] Stuart T. Johnston, Matthew J. Simpson, and Michael J. Plank. Lattice-free descriptions of collective motion with crowding and adhesion. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 88(6):062720, 2013.

[59] Michael J Plank and Matthew J Simpson. Models of collective cell behaviour with crowding effects: comparing lattice-based and lattice-free approaches. *Journal of the Royal Society, Interface / the Royal Society*, 9(76):2983–96, 2012.

[60] B. Szabo, G. J. Szöllösi, B. Gönci, Zs Juranyi, D. Selmeczi, and Tamas Vicsek. Phase transition in the collective migration of tissue cells: Experiment and model. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 74(6):1–5, 2006.

[61] Nestor Sepulveda, Laurence Petitjean, Olivier Cochet, Erwan Grasland-Mongrain, Pascal Silberzan, and Vincent Hakim. Collective Cell Motion in an Epithelial Sheet Can Be Quantitatively Described by a Stochastic Interacting Particle Model. *PLoS Computational Biology*, 9(3), 2013.

[62] M Deforet, V Hakim, H G Yevick, G Duclos, and P Silberzan. Emergence of collective modes and tri-dimensional structures from epithelial confinement. *Nature communications*, 5(May):3747, 2014.

[63] Liang Li, Edward C Cox, and Henrik Flyvbjerg. 'Dicty dynamics': Dictyostelium motility as persistent random motion. *Physical biology*, 8(4):046006, 2011.

[64] Claus Metzner, Christoph Mark, Julian Steinwachs, Lena Lautscham, Franz Stadler, and Ben Fabry. Superstatistical analysis and modelling of heterogeneous random walks. *Nature Communications*, 6(May):7516, 2015.

[65] Edward A Codling, Michael J Plank, and Simon Benhamou. Random walk models in biology. *Journal of the Royal Society, Interface / the Royal Society*, 5(25):813–834, 2008.

[66] M H Gail and C W Boone. The locomotion of mouse fibroblasts in tissue culture. *Biophysical journal*, 10(10):980–93, 1970.

[67] R S Hartman, K Lau, W Chou, and T D Coates. The fundamental motor of the human neutrophil is not random: evidence for local non-Markov movement in neutrophils. *Biophysical journal*, 67(6):2535–2545, 1994.

[68] David Selmeczi, Stephan Mosler, Peter H Hagedorn, Niels B Larsen, and Henrik Flyvbjerg. Cell motility as persistent random motion: theories from experiments. *Biophysical journal*, 89(2):912–931, 2005.

[69] Peter Dieterich, Rainer Klages, Roland Preuss, and Albrecht Schwab. Anomalous dynamics of cell migration. *Proceedings of the National Academy of Sciences*, 105(2):20, 2009.

[70] Liang Li, Simon F. Nørrelkke, and Edward C. Cox. Persistent cell motion in the absence of external signals: A search strategy for eukaryotic cells. *PLoS ONE*, 3(5), 2008.

[71] Alka A. Potdar, Jenny Lu, Junhwan Jeon, Alissa M. Weaver, and Peter T. Cummings. Bimodal analysis of mammary epithelial cell migration in two dimensions. *Annals of Biomedical Engineering*, 37(1):230–245, 2009.

[72] H.C. Berg. *E. coli in motion*. 2004.

[73] Alka A. Potdar, Junhwan Jeon, Alissa M. Weaver, Vito Quaranta, and Peter T. Cummings. Human mammary epithelial cells exhibit a bimodal correlated random walk pattern. *PLoS ONE*, 5(3), 2010.

[74] Delphine Arcizet, Sofia Capito, Mari Gorelashvili, Carolin Leonhardt, Marion Vollmer, Simon Youssef, Susanne Rappl, and Doris Heinrich. Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces. *Soft Matter*, 8(5):1473, 2012.

[75] Mari Gorelashvili, Martin Emmert, Kai F. Hodeck, and Doris Heinrich. Amoeboid migration mode adaption in quasi-3D spatial density gradients of varying lattice geometry. *New Journal of Physics*, 16, 2014.

[76] F. Detcheverry. Unimodal and bimodal random motions of independent exponential steps. *European Physical Journal E*, 37(11):1–9, 2014.

[77] F. Detcheverry. Non-Poissonian run-and-turn motions. *EPL (Europhysics Letters)*, 111(6):60002, 2015.

[78] Fernando Peruani and Luis G. Morelli. Self-propelled particles with fluctuating speed and direction of motion in two dimensions. *Physical Review Letters*, 99(1):1–4, 2007.

[79] Laurent Golé. Migration of Dictyostelium amoeba: role of adhesion and quorum-sensing. *Thèse de doctorat*, 2011.

[80] Gabriel Rosser, Alexander G. Fletcher, David A. Wilkinson, Jennifer A. de Beyer, Christian A. Yates, Judith P. Armitage, Philip K. Maini, and Ruth E. Baker. Novel Methods for Analysing Bacterial Tracks Reveal Persistence in Rhodobacter sphaeroides. *PLoS Computational Biology*, 9(10), 2013.

[81] Steve Atkinson and Paul Williams. Quorum sensing and social networking in the microbial world. *Journal of the Royal Society, Interface / the Royal Society*, 6(40):959–978, 2009.

[82] Michael Kruppa. Quorum sensing and Candida albicans. *Mycoses*, 52(1):1–10, 2009.

[83] Alexandra Kolbinger, Tong Gao, Debbie Brock, Axel Kisters, Joseph Kellermann, Diane Hatton, Richard H Gomer, Birgit Wetterauer, and Robin Ammann. A Cysteine-Rich Extracellular Protein Containing a PA14 Domain Mediates Quorum Sensing in Dictyostelium discoideum A Cysteine-Rich Extracellular Protein Containing a PA14 Domain Mediates Quorum Sensing in Dictyostelium discoideum. *Eukaryotic cell*, 4(6):991–998, 2005.

[84] M. Clarke and R. H. Gomer. PSF and CMF, autocrine factors that regulate gene expression during growth and early development of Dictyostelium. *Experientia*, 51(12):1124–1134, 1995.

[85] Pauline Schaap. Evolution of developmental cyclic adenosine monophosphate signaling in the Dictyostelia from an amoebozoan stress response. *Development Growth and Differentiation*, 53(4):452–462, 2011.

[86] G M Souza, a M da Silva, and a Kuspa. Starvation promotes Dictyostelium development by relieving PufA inhibition of PKA translation through the YakA kinase pathway. *Development (Cambridge, England)*, 126(14):3263–3274, 1999.

[87] R. Jain, I. S. Yuen, C. R. Taphouse, and R. H. Gomer. A density-sensing factor controls development in Dictyostelium. *Genes and Development*, 6(3):390–400, 1992.

[88] Lei Tang, Robin Ammann, Tong Gao, and Richard H. Gomer. A Cell Number-counting Factor Regulates Group Size in Dictyostelium by Differentially Modulating cAMP-induced cAMP and cGMP Pulse Sizes. *Journal of Biological Chemistry*, 276(29):27663–27669, 2001.

[89] Lei Tang, Tong Gao, Catherine McCollum, Wonhee Jang, Michael G Vicker, Robin R Ammann, and Richard H Gomer. A cell number-counting factor regulates the cytoskeleton and cell motility in Dictyostelium. *Proceedings of the National Academy of Sciences of the United States of America*, 99(3):1371–6, 2002.

[90] Najl V Valeyev, Jung-Su Kim, J S Pat Heslop-Harrison, Ian Postlethwaite, Nicolay V Kotov, and Declan G Bates. Computational modelling suggests dynamic interactions between Ca2+, IP3 and G protein-coupled modules are key to robust Dictyostelium aggregation. *Molecular bioSystems*, 5(6):612–628, 2009.

[91] Gareth Bloomfield and Catherine Pears. Superoxide signalling required for multicellular development of Dictyostelium. *Journal of cell science*, 116(Pt 16):3387–3397, 2003.

[92] JA Hadwiger, S Lee, and RA Firtel. The G alpha subunit G alpha 4 couples to pterin receptors and identifies a signaling pathway that is essential for multicellular development in Dictyostelium. *Proceedings of the National Academy of Sciences of the United States of America*, 91(22):10566–10570, 1994.

[93] Hoai Nghia Nguyen and Jeffrey A. Hadwiger. The Gα4 G protein subunit interacts with the MAP kinase ERK2 using a D-motif that regulates developmental morphogenesis in Dictyostelium. *Developmental Biology*, 335(2):385–395, 2009.

[94] B.V. Hinkley. On the ratio of two corelated normal random variables. *Biometrika*, 56:635–639, 1969.

[95] S J Pirt. A kinetic study of the mode of growth of surface colonies of bacteria and fungi. *Journal of General Microbiology*, pages 181–197, 1967.

[96] S A Palumbo, M G Johnson, V T Rieck, and L D Witter. Growth measurements on surface colonies of bacteria. *Journal of general microbiology*, 66(2):137–143, 1971.

[97] J.W.T. Wimpenny. The Growth and Form of Bacterial Colonies. *Journal of general microbiology*, 114:483–486, 1979.

[98] J D Murray. *Mathematical Biology : I . An Introduction , Third Edition*, volume 1. 2002.

[99] R. A. Fisher. the Wave of Advance of Advantageous Genes. *Annals of Human Genetics*, 7:355–369, 1937.

[100] A.N. Kolmogorov, I.G. Petrovskii, and N.S. Piskunov. Etude de l'équation de la diffusion avec croissance de la quantité de matiere et son applicationa un probleme biologiqueNo Title. *Bulletin Mathématique de l'Université de Moscou*, 1937.

[101] J.D. Murray. Mathematical Biology I: An Introduction. 2001.

[102] M. Block, E. Schöll, and D. Drasdo. Classifying the expansion kinetics and critical surface dynamics of growing cell populations. *Physical Review Letters*, 99(24):3–6, 2007.

[103] Helen Byrne and Dirk Drasdo. Individual-based and continuum models of growing cell populations: A comparison. *Journal of Mathematical Biology*, 58(4-5):657–687, 2009.

[104] Antonio Brú, Juan Pastor, Isabel Fernaud, Isabel Brú, Sonia Melle, and Carolina Berenguer. Super-Rough Dynamics on Tumor Growth. *Physical Review Letters*, 81(18):4008–4011, 1998.

[105] Antonio Brú, Sonia Albertos, José Luis Subiza, José López García-Asenjo, and Isabel Brú. The universal dynamics of tumor growth. *Biophysical journal*, 85(November):2948–2961, 2003.

[106] J Galeano, J Buceta, K Juarez, B Pumariño, J. De La Torre, and J. M Iriondo. Dynamical scaling analysis of plant callus growth. *Europhysics Letters (EPL)*, 63(July):83–89, 2007.

[107] M.A.C. Huergo, M. A. Pasquale, P. H. González, A. E. Bolzán, and A. J. Arvia. Dynamics and morphology characteristics of cell colonies with radially spreading growth fronts. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 84(2):1–11, 2011.

[108] M. A C Huergo, M. A. Pasquale, P. H. González, A. E. Bolzán, and A. J. Arvia. Growth dynamics of cancer cell colonies and their comparison with noncancerous cells. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85(1):1–9, 2012.

[109] Bram G Sengers, Colin P Please, and Richard O C Oreffo. Experimental characterization and computational modelling of two-dimensional cell spreading for skeletal regeneration. *Journal of the Royal Society, Interface / the Royal Society*, 4(17):1107–1117, 2007.

[110] Matthew J Simpson, Katrina K Treloar, Benjamin J Binder, Parvathi Haridas, Kerry J Manton, David I Leavesley, D L Sean McElwain, and Ruth E Baker. Quantifying the roles of cell motility and cell proliferation in a circular barrier assay Quantifying the roles of cell motility and cell proliferation in a circular barrier assay. *Journal of the Royal Society, Interface / the Royal Society*, 2013.

[111] Anna-Kristina Marel, Matthias Zorn, Christoph Klingner, Roland Wedlich-Söldner, Erwin Frey, and Joachim O. Rädler. Flow and Diffusion in Channel-Guided Cell Migration. *Biophysical Journal*, 107(5):1054–1064, 2014.

[112] Massimiliano Maria Baraldi, Alexander a Alemi, James P Sethna, Sergio Caracciolo, Caterina a M La Porta, and Stefano Zapperi. Growth and form of melanoma cell colonies. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(02):P02032, 2013.

[113] M Poujade, E Grasland-Mongrain, a Hertzog, J Jouanneau, P Chavrier, B Ladoux, a Buguin, and P Silberzan. Collective migration of an epithelial monolayer in response to a model wound. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41):15988–15993, 2007.

[114] Frederick K Balagaddé, Lingchong You, Carl L Hansen, Frances H Arnold, and Stephen R Quake. Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science (New York, N.Y.)*, 309(2005):137–140, 2005.

[115] John H. Koschwanez, Robert H. Carlson, and Deirdre R. Meldrum. Thin PDMS films using long spin times or tert-butyl alcohol as a solvent. *PLoS ONE*, 4(2):2–6, 2009.

[116] Maria Bruna and S. Jonathan Chapman. Diffusion of multiple species with excluded-volume effects. *Journal of Chemical Physics*, 137(20), 2012.

[117] M E Cates and J Tailleur. When are active Brownian particles and run-and-tumble particles equivalent? Consequences for motility-induced phase separation. *Europhysics Letters*, 101(2):20010, 2013.

[118] P Friedl, S Borgmann, and E B Bröcker. Amoeboid leukocyte crawling through extracellular matrix: lessons from the Dictyostelium paradigm of cell movement. *Journal of leukocyte biology*, 70(4):491–509, 2001.

[119] Peter Friedl and Katarina Wolf. Plasticity of cell migration: A multiscale tuning model. *Journal of Cell Biology*, 188(1):11–19, 2010.

[120] Silke Henkes, Yaouen Fily, and M. Cristina Marchetti. Active jamming: Self-propelled soft particles at high density. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 84(4):84–87, 2011.

[121] Jin-Ah Park, Jae Hun Kim, Dapeng Bi, Jennifer A. Mitchel, Nader Taheri Qazvini, Kelan Tantisira, Chan Young Park, Maureen McGill, Sae-Hoon Kim, Bomi Gweon, Jacob Notbohm, Robert Steward Jr, Stephanie Burger, Scott H. Randell, Alvin T. Kho, Dhananjay T. Tambe, Corey Hardin, Stephanie A. Shore, Elliot Israel, David A. Weitz, Daniel J. Tschumperlin, Elizabeth P. Henske, Scott T. Weiss, M. Lisa Manning, James P. Butler, Jeffrey M. Drazen, and Jeffrey J. Fredberg. Unjamming and cell shape in the asthmatic airway epithelium. *Nature Materials*, 14(10):1040–1048, 2015.

[122] Haiyang Zhang, Stuart Angus, Michael Tran, Chunyan Xie, Oleg A. Igoshin, and Roy D. Welch. Quantifying aggregation dynamics during Myxococcus xanthus development. *Journal of Bacteriology*, 193(19):5164–5170, 2011.

[123] Darja Dubravcic, Minus van Baalen, and Clément Nizak. An evolutionarily significant unicellular strategy in response to starvation stress in Dictyostelium social amoebae. *F1000Research*, 3(0):133, 2014.

[124] K Müller and G Gerisch. A specific glycoprotein as the target site of adhesion blocking Fab in aggregating Dictyostelium cells. *Nature*, 274(5670):445–9, 1978.

[125] F. D C Farrell, M. C. Marchetti, D. Marenduzzo, and J. Tailleur. Pattern formation in self-propelled particles with density-dependent motility. *Physical Review Letters*, 108(24):1–5, 2012.

[126] P. Friedl and K. Wolf. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nature reviews. Cancer*, 3(5):362–74, 2003.

Contents

Abstract

Cell motility is fundamental in many physiological, either normal or pathological, phenomena. Yet, although these most often involve several cells moving at the same time, how the interactions between cells affect both individual and collective dynamics remains a poorly understood question. In this thesis, I used vegetative *Dictyostelium discoideum* cells as a model to study this collective regulation of the motility. I relied mainly on the thorough analysis of numerous cell trajectories in various situations to (*i*) characterise a secreted factor used to down-regulate the cells' motility (biochemical nature, response pathway, secretion and response dynamics) and (*ii*) quantitatively analyse and model the dynamics of spreading cell colonies of controlled initial shape, size and density. Last, I describe a dynamic aggregation phenomenon that occurs when the cells are seeded at high density in a nutrient-rich medium.

Résumé

La motilité cellulaire est fondamentale dans de nombreux processus physiologiques, qu'ils soient normaux ou pathologiques. Cependant, bien que ces derniers impliquent la plupart du temps de nombreuses cellules se mouvant en même temps, les effets des interactions entre cellules sur leur dynamique, à la fois individuelle et collective, restent assez mal connu. Dans cette thèse, j'ai utilisé *Dictyostelium discoideum* à l'état végétatif pour étudier cette régulation collective de la motilité. Je me suis principalement appuyé sur une analyse minutieuse de nombreuses trajectoires cellulaires dans des situations variées pour (*i*) caractériser un facteur sécrété qui régule négativement la motilité cellulaire (nature chimique, voie de signalisation, dynamique de sécrétion et de réponse) et (*ii*) analyser et modéliser quantitativement la dynamique d'étalement de colonie cellulaires de forme, dimension et densité contrôlées. Je décris enfin un phénomène d'agrégation dynamique observé lorsque les cellules sont placées à haute densité dans un milieu nutritif.