High fidelity information processing in folic acid chemotaxis of Dictyostelium amoebae

Igor Segota, Surin Mong, Eitan Neidich, Archanachandu, Catherine J. Lussenhop and Carl Franck

Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY 14853 (USA)

[accepted to Journal of the Royal Society Interface]

Living cells depend upon the detection of chemical signals for their existence. Eukaryotic cells can sense a concentration difference as low as a few percent across their bodies. This process was previously suggested to be limited by the receptor-ligand binding fluctuations. Here, we first determine the chemotaxis response of Dictyostelium cells to static folic acid gradients and show that they can significantly exceed this sensitivity, responding to gradients as shallow as 0.2% across the cell body. Second, using a previously developed information theory framework, we compare the total information gained about the gradient (based on the cell response) to its upper limit: the information gained at the receptor-ligand binding step. We find that the model originally applied to cAMP sensing fails as demonstrated by the violation of the data processing inequality, i.e. the total information exceeds the information at the receptor-ligand binding step. We propose an extended model with multiple known receptor types and with cells allowed to perform several independent measurements of receptor occupancy. This does not violate the data processing inequality and implies the receptor-ligand binding noise dominates both for low and high chemoattractant concentrations. We also speculate that the interplay between exploration and exploitation is employed as a strategy for accurate sensing of otherwise unmeasurable levels of a chemoattractant.

INTRODUCTION

Eukaryotic amoebae Dictyostelium discoideum (referred as Dictyostelium) in vegetative state forage on bacteria by following gradients of folic acid, a by-product of bacterial metabolism [1][2]. It is currently believed that Dictyostelium measure chemical gradients directly by monitoring the distribution of the occupied chemoattractant receptors. These cells can detect concentration differences as low as a few percent across their cell bodies [3][8] and it is currently an open question what exactly limits this process. Previously, the receptor-ligand binding fluctuations were suggested as the limiting factor, which remains a possibility since a single excited receptor may amplify the signal by activating multiple G-proteins [9][11].

Chemotaxis signaling system can be described as the following Shannon communication channel [12][13]: the chemoattractant gradient direction as the input, the spatial distribution of occupied receptors as the intermediate step and the direction of cell motion as the output. Fuller et al. [4] recently exploited this information-theoretic framework, where a cell in a static gradient was modelled as N receptors arranged in a circle, each in chemical equilibrium with the local chemoattractant concentration, described by a dissociation constant $K_d$.

The joint state of all receptors $\theta_{res}$ was assumed to depend only on the gradient direction, $\theta_{grad}$. Likewise, the probability of cell moving in a direction $\theta_{res}$ was assumed to depend only on $\theta_{rec}$, with these three variables forming a Markov chain: $\Theta_{grad} \rightarrow \Theta_{rec} \rightarrow \Theta_{res}$ (see SI). Capital letters denote random variables and lowercase their values. Fuller et al. [4] computed the mutual information between the gradient direction and the receptor distribution $I_{ext}(\theta_{grad}, \theta_{rec})$ as “external mutual information”, $I_{ext}$ quantifies the information gained about the gradient through a perfect (noiseless) readout of the occupied receptors.

Furthermore, Fuller et al. [4] used Dictyostelium cAMP chemotaxis experiments to calculate the mutual information between the gradient direction and the cell response $I_{tot}(\Theta_{grad}, \Theta_{res})$, “total mutual information”. $I_{tot}$ quantifies the information gained about the gradient by cells through the imperfect (noisy) readout of the occupied receptors. The data processing inequality [14] (p.34) states that in a Markov chain of variables, the information can only be destroyed in each subsequent step, which here translates into $I_{tot} \leq I_{ext}$. In other words, the information gained by cells after being processed through the entire signaling pathway, cannot exceed the information gained at the receptor level. The authors [4] then argued that for low cAMP concentrations the receptor-ligand binding fluctuations dominate the entire noise ($I_{tot} \approx I_{ext}$), since there is no further information loss downstream. Previously, Ueda and Shibata [11] also reached this conclusion using signal-to-noise ratio arguments, using stochastic receptor noise and time integration with second messengers and locomotion systems.

Here, we measure the response of a population of Dictyostelium cells to static linear FA gradients, established in an agarose-gel based microfluidic device [15]. The steady state gradients were achieved by maintaining fixed concentrations of FA on opposite sides of a microfluidic channel (see SI). A linear gradient was established by diffusion through agarose gel. Cell migration was recorded using time-lapse optical microscopy. The measured distribution of cell displacement angles $p(\theta_{res} | \theta_{grad})$ was used to calculate the total mutual information $I_{tot}$ and compared to $I_{ext}$ (using the result in [4]) to test the possibility of receptor-ligand binding fluctuations dominating the total noise.

RESULTS AND DISCUSSION

First, we employ the result in Fuller et al. [4], Eq.S56 for the external mutual information $I_{ext}$ for shallow linear
where \( c(x) \) is the concentration measured in units of \( K_d \), \( \nabla c \) is the gradient measured in units of \( K_d/R \) (\( R \) is the radius of a hemispherical cell, taken as \( 5\mu\text{m} \)) and the dimensionless small parameter \( \epsilon \equiv \nabla c/(1 + c(x)) \ll 1 \). For larger values of \( \epsilon \) one has to resort to numerical simulations. The design of our microfluidic device ensured it was applicable to use the Eq.\( \ref{eq:1} \) as the small parameter was in range \( 0.0003 \leq \epsilon \leq 0.0065 \).

Previously, Wurster and Butz \( \cite{16} \) and de Wit and van Haastert \( \cite{17} \) measured the dissociation constants \( K_d \) and receptor numbers \( N \) using radioligand assays. In the former case \( \cite{16} \), we used the measured \( N \) and \( K_d \) after 3 hours in the buffer, which reflects the conditions in our experiments. In the latter case \( \cite{17} \) vegetative cells were employed. Wurster and Butz \( \cite{16} \) found \( K_d = 150 \text{nM} \), \( N = 60,000 \) and de Wit and van Haastert \( \cite{17} \) found five receptor types with the following dissociation constants and receptor numbers: 1) \( K_d1 = 450 \text{nM} \), \( N_1 = 80,000 \), 2) \( K_d2 = 70 \text{nM} \), \( N_2 = 80,000 \), 3) \( K_d3 = 17 \text{nM} \), \( N_3 = 550 \), 4) \( K_d4 = 50 \text{nM} \), \( N_4 = 50 \) and 5) \( K_d5 = 15 \text{nM} \), \( N_5 = 1,450 \). In both cases Scatchard plots show that the first-order kinetics can be employed with good approximation but that there is slight curvature implying either negative cooperativity or greater receptor heterogeneity. Furthermore, the binding curves for folic acid were taken up to \( \mu\text{M} \) concentrations, the interesting range explored in this study.

Second, we measured the cell trajectories and the distribution of angles \( p(\theta_{res}|\theta_{grad}) \) of total displacement vectors (Fig.\( \ref{fig:1}a \)) of a population of Dictyostelium cells (see SI for Methods). In each experiment the FA gradient was uniform and the concentration varied at most threefold across the width of a channel. Each experiment was repeated until we obtained 300 to 700 cell trajectories. These experiments were used to calculate the total mutual information \( I_{tot} \) and the chemotactic index (CI). CI is defined as  

\[
CI = (\sum_i \bar{r}_{i}) \cdot \bar{n} / \left( \sum_i |\bar{r}_{i}| \right)
\]

where \( \bar{r}_{i} \) is the instantaneous cell displacement during the time step \( i \) (taken as 30 seconds) and \( \bar{n} \) is the gradient direction.

We performed ten experiments where we varied the FA concentration in the top channel of microfluidic device while keeping the bottom channel at concentration zero. In these experiments both the concentration and the gradient were changed and these are plotted in Fig.\( \ref{fig:2}a \). We also performed five additional experiments (shown in Fig.\( \ref{fig:2}a \) and \( \ref{fig:2}b \)) where we changed the mean concentration and the gradient separately. For the range of concentrations and gradients explored here, decreasing the gradient and increasing FA concentration diminished the signal. Therefore, the FA chemotaxis can depend both on the absolute value of FA concentration and its gradient.

\[ I_{tot} = \frac{N}{4 \ln 2} \left( \frac{\nabla c}{1 + c(x)} \right)^2 \]  

(1)

where \( c(x) \) is the concentration measured in units of \( K_d \), \( \nabla c \) is the gradient measured in units of \( K_d/R \) (\( R \) is the radius of a hemispherical cell, taken as \( 5\mu\text{m} \)) and the dimensionless small parameter \( \epsilon \equiv \nabla c/(1 + c(x)) \ll 1 \). For larger values of \( \epsilon \) one has to resort to numerical simulations. The design of our microfluidic device ensured it was applicable to use the Eq.\( \ref{eq:1} \) as the small parameter was in range \( 0.0003 \leq \epsilon \leq 0.0065 \).

FIG. 1. Measured chemotaxis response for a range of gradients and mean concentrations. (a) Distribution of cell displacement angles for the peak response for the gradient \( dc/dx = 1.6 \text{nM}/\mu\text{m} \) and mean concentration \( \text{c}_0 = 2,500 \text{nM} \). Each radial step represents 15 data points. (b) CI for experiments with variable FA concentration in the top channel, which changed both the mean concentration and the gradient. The controls denote CI for experiments performed with no gradient with mean FA concentrations of 0.2,500 nM and 10,000 nM. The error bars and gray area denote Standard Error of the Mean (SEM).

\[ I_{tot} = \frac{N}{4 \ln 2} \left( \frac{\nabla c}{1 + c(x)} \right)^2 \]  

(1)

with the error due to finite number of data points estimated as \((m - 1)/2M \) \( \cite{15} \), where \( M \) is the total number of data points.

Next, we compare \( I_{tot} \) and \( I_{ext} \). Fig.\( \ref{fig:2}a \) shows that for low concentrations and shallow gradients \( I_{tot} \approx I_{ext} \), meaning the receptor-ligand binding fluctuations dominate the total noise. This possibility was previously suggested for cAMP \( \cite{12} \) using signal-to-noise ratio analysis with a biased random walk model of cell motion. The information-theoretic analysis assumes only the steady-state receptor-ligand binding fluctuations and benefits from not being tied to a particular model of cell motion, since Dictyostelium cells do not follow a simple random walk \( \cite{19} \).

The most surprising result is that the response is observed for gradients as low as 0.2% across the cell body \((dc/dx = 3.2 \text{nM}/\mu\text{m}) \), \( \text{c}_0 = 15,000 \text{nM} \), \( I_{tot} = 0.06 \) bits shown in Fig.\( \ref{fig:2}b \). For these experiments, the difference in the fraction of occupied receptors front-to-back on the cell body is given by:

\[
\eta = \frac{c_{front}}{c_{front} + K_d} - \frac{c_{back}}{c_{back} + K_d}
\]  

(3)

and is shown in Table 1 for different measured dissociation constants. This fraction is at most 0.006% which amounts to a 1-10 receptors difference with 29,700 receptors (or 99%) occupied on each side, indicating a highly saturated regime. Furthermore, in this range the data processing inequality \((I_{tot} \leq I_{ext})\) is strongly violated as we have \( I_{tot} > I_{ext} \). The observed response is better than theoretically possible with receptor-ligand binding fluctuations as the only noise

\[
\eta = \frac{c_{front}}{c_{front} + K_d} - \frac{c_{back}}{c_{back} + K_d}
\]  

(3)
FIG. 2. Comparison of the total mutual information $I_{\text{tot}}$ and external mutual information $I_{\text{ext}}$. (a) $I_{\text{tot}}$ (dashed line) and $I_{\text{ext}}$ (solid line) for the same experiments as in Fig. 1b, both averaged over all local concentrations (see SI). Error bars for $I_{\text{tot}}$ represent SEM. The shaded range for $I_{\text{ext}}$ denotes its spread due to the range of local concentrations the cells were exposed to in the microfluidic device (see text). Dotted line denotes the $I_{\text{tot}}$ for control experiments without a gradient. Annotation 1 shows the range where the data processing inequality is strongly violated, $I_{\text{tot}} > I_{\text{ext}}$. (b) Left: calculated values for $I_{\text{ext}}$ (Eq.1); shaded area denotes the combinations of $c_0$ and $dc/dx$ inaccessible in our experiment due to the geometry of the microfluidic device and low solubility of FA in DB ($\sim 0.1$ mM). Right: the range of concentrations and gradients where cAMP chemotaxis has been measured, colored by the value of measured chemotactic index, CI. The measurement with annotation 1 ($c_0 = 500$ nM = 17 $K_d$, $dc/dx = 0.5$ nM/μm = 0.08 $K_d/R$, CI=0.25) is done in approximate range where we detected the greatest violation of the data processing inequality ($c_0 = 5,000$ nM = 33 $K_d$, $dc/dx = 3.2$ nM/μm = 0.11 $K_d/R$, CI=0.13, $I_{\text{tot}} = 0.16$ bits) if we compare them by rescaling the concentrations with their respective $K_d$ ($K_d$(cAMP) = 30 nM, $K_d$(folic acid) = 150 nM [10, 20]). $I_{\text{tot}}$ and $I_{\text{ext}}$ for experiments with fixed gradient, where mean concentration is changed is shown in (c) and experiments with fixed mean concentration where the gradient is changed are shown in (d). In the range investigated here, increasing the concentration and reducing the gradient reduced the chemotaxis response, $I_{\text{tot}}$ but the violation of the data processing inequality persists.

source. Next, we compared our results with previous cAMP chemotaxis experiments [3, 5, 7, 10, 11, 12], shown in Fig 2. In comparing critical parameters, the receptor-ligand binding constant $K_d$(FA) = 150 nM stands out as a factor of five greater compared to cAMP, $K_d$(cAMP) = 30 nM, whereas the number of receptors per cell is almost the same: 60,000 for FA and 70,000 for cAMP [16, 20].

$$
\begin{array}{c|ccccc}
K_d \text{nM} & 450 & 150 & 70 & 50 & 17 & 15 \\
\eta & 0.006 \% & 0.002 \% & 0.001 \% & 0.0007 \% & 0.0002 \% & 0.0002 \%
\end{array}
$$

TABLE I. Fraction of occupied receptors front to back of the cell for the shallowest gradient where we measured the chemotaxis response, calculated using each measured receptor type according to Eq.3.

These simplified descriptions of FA and cAMP receptors were sufficient to explain the results in [1], but do not suffice here – possibly explained by the limited range of cAMP concentrations and gradients investigated in [1] (see Fig 2). The measurement with annotation 1 on Fig 2, from Var-
Effects of multiple receptor types and receptor phosphorylation

Second, we considered all different receptor types mentioned previously. This possibility was motivated by the local minimum in $I_{\text{tot}}$ shown in Fig.2a indicating that perhaps there are two receptor types or states, each active in a distinct range of local ligand concentrations. We calculated $I_{\text{ext}}$ for each receptor type and added it together to investigate whether this resolves the violation of the data processing inequality. The results are shown in Fig.3a and indicate that the presence of multiple receptor types reduces, but does not eliminate the violation of the data processing inequality. This is because the shaded range for $I_{\text{ext}}$ in Fig.3 represents the range of concentrations the cells were exposed to in our microfluidic device (and not the uncertainty), with the maximum value of $I_{\text{ext}}$ corresponding to the bottom of our device and the minimum value corresponds to the top of our device. However, the systematic uncertainty of the average $I_{\text{ext}}$ (solid line in Fig.3) is only 10% (see SI), which is what is compared to the average $I_{\text{tot}}$ [22]. Furthermore, the double-peak feature observed in $I_{\text{tot}}$ is not exactly reproduced in $I_{\text{ext}}$ even when considering only two receptor types. This could be due to the fact that $I_{\text{ext}}$ is only an upper limit for $I_{\text{tot}}$ and in this range the intracellular signal processing is not negligible, so $I_{\text{tot}} < I_{\text{ext}}$. In other words the dip could be the consequence of the extra noise somewhere downstream of the receptor-ligand binding events. It is also worth mentioning that this double-peak response prevents us from using any single receptor with fixed $K_d$ to fix the violation of the data processing inequality, unless the receptor number per cell $N$ is set to a factor 12 more than it is measured.

Therefore, this explanation could be plausible only if all the cells were concentrated near the bottom of our device. In our experiments they were always uniformly distributed with the mean position in the center.

Third, we hypothesized that FA receptors can be phosphorylated. Xiao et al. [23] have shown that the phosphorylation of cAMP receptors cAR1 reduced the affinity (increased $K_d$) of a cAMP-cAR1 process by a factor of three, from 300 nM to 900 nM. Here we assume that the additional receptor types can be phosphorylated to $3 \times K_{\text{ext}}$ and fit the data in the same way as for additional receptor types. The results (Fig.4) show that this only reduced the violation of the data processing inequality, but did not eliminate it.

Effects of cell polarization

Fourth, we considered for the possibility of cell polarization [24], previously considered in [13, 25]. In our analysis thus far, we assumed that cells had no previous knowledge of the gradient direction, so the prior probability was $p(\theta_{\text{grad}}) = 1/2\pi$. Now, we consider a circular normal prior distribution:

$$p(\theta_{\text{grad}}) = \frac{\exp(K \cos \theta_{\text{grad}})}{2\pi I_0(K)}$$

where $I_0(K)$ is the modified Bessel function of the first kind of zeroth order, and the parameter $K$ measures the bias strength. We used the approach from [24] to numerically calculate $I_{\text{bias}}^{\text{tot}}(K)$. We also numerically calculated $I_{\text{bias}}^{\text{tot}}(K)$ (see SI) and then compared both $I_{\text{bias}}^{\text{tot}}$ and $I_{\text{bias}}^{\text{ext}}$ up to very biased distributions with $K = 80$, as larger values required significantly higher numerical precision. Fig.3 shows that the violation of the data processing inequality still persists.

Effects of multiple measurements

Finally we investigated the effect of multiple independent measurements of the receptor occupancy [7, 11], occurring if cells can choose between i) short and imprecise gradient measurements, but moving fast and ii) long and precise gradient measurements but moving more slowly. This is known as the tradeoff between exploration and exploitation in the field of reinforcement learning [20].

Eq.1 is only valid for a single snapshot measurement. The information acquired from multiple independent measurements is simply the sum of the information of each contribution due to a single measurement. Therefore we multiply the Eq.1 by the number of independent measurements $N_{\text{meas}} = T_{\text{pseudo}}/T_{\text{correl}}$ [4, 25], where $T_{\text{pseudo}}$ is the time scale of pseudopod extension and $T_{\text{correl}}$ is the receptor correlation time (this ratio gives us the maximum number of measurements that could have been performed). We note that $T_{\text{pseudo}}$ is likely the upper bound for the integration time based on the evidence in variable gradient experiments [27] where it was observed that the cells extend their pseudopods in the gradient direction as soon as the direction of the gradient is changed. Rappel and Levine previously noted that the correlation time consists of both receptor chemical dynamics and the diffusive process and estimated the cAMP receptor correlation time $T_{\text{correl}} = 5s$ [28, 29]. Fuller et al. [4] concluded $N_{\text{meas}} \approx 1$. We estimated $N_{\text{meas}}^{\text{FA}}$ by assuming that $T_{\text{pseudo}}$ is inversely proportional to the mean cell speed, and the same $T_{\text{correl}}$ for both FA and cAMP receptors (based on comparable receptor off-rates for FA and cAMP receptors [17, 30]):

$$N_{\text{meas}}^{\text{FA}} = \frac{T_{\text{pseudo}}^{\text{FA}}}{T_{\text{correl}}^{\text{FA}}} \approx N_{\text{meas}}^{\text{cAMP}} \frac{v_{\text{cAMP}}}{v_{\text{FA}}}$$

where the chemotaxis speeds are: $v_{\text{cAMP}} = 0.25 \mu m/s$ [4] and $0.05 \mu m/s \leq v_{\text{FA}} \leq 0.12 \mu m/s$, which gives $2 \leq N_{\text{meas}}^{\text{FA}} \leq 4$. $I_{\text{tot}}$ and $I_{\text{ext}}$ are compared in Fig.3 and show that this only reduced the violation of the data processing inequality, but again does not eliminate it. Recently developed approaches considered diffusible inhibitors in balanced inactivation model [28, 29, 31] and their integration time ($T_{\text{int}} = 10 s$) corresponds roughly to the integration times estimated here ($T_{\text{int}} = 10$ to 20 s). In addition, the models considered so far do not reproduce the double peak observed experimentally (Fig.3), but this might be the consequence of a significant information loss downstream of the receptor-ligand binding events.
However, combining the effects of additional receptor types and multiple independent measurements does not result in the violation of the data processing inequality; see Fig.3b. $N^{FA}_{meas}$ roughly agrees with $N^{cAMP}_{meas} \approx 2$, which was included to explain a much greater range of concentrations and gradients than in [31] (see Fig.2).

It should still be noted that the multiple independent measurements can be a consequence of integrating the information from multiple pseudopods [24]. During the 30 second time interval cells extend a number of small protrusions (sometimes simultaneously), some of which are retracted quickly (see Fig.6 in [24]). Taking this into account would lead to a different definition of the total mutual information than that used here where the centroid of each cell is used to specify its position. One direction for future studies is then to perform experiments with higher resolution to quantify the information acquired about the gradient, employing this alternative measure.

**Other effects**

Fig.2 implies that the total noise is indeed dominated by the receptor-ligand binding fluctuations at both low and high gradients and concentrations. This seems plausible since in that range the receptors are either mostly unoccupied or occupied. In the intermediate range where $I_{ext} \gg I_{tot}$, the internal noise dominates. We note that it has been shown [33] that there is always a fraction of cell population which does not respond to gradients and polarizes in random directions, independent of the external cAMP gradient. Since in our experiments we only have static gradients we could not separately identify these cells and they had to be included in the data analysis. Exclusion of this subpopulation from our analysis would increase the total mutual information $I_{tot}$ even further and the violation of the data processing inequality would be even larger.

The possibility of receptor interactions was ruled out due to uniform receptor distributions for both FA [34] and cAMP receptors [35, 36]. Unlike in the cAMP case [4], here the non-circularity of cell shapes is not an issue since the cells are circular when sensing FA. However, there is still a possibility of a more complicated mechanism if FA receptors also transport FA into the cell [9], serving as a different communication channel, or if a FA transporter is a separate protein, as a separate communication channel.

The possibility that FA and cAMP receptors share the majority of the internal signaling pathway [32] implies equal FA and cAMP responses, if rescaled by their respective parameters $K_d$ and $N$. This remains to be investigated with more cAMP and FA chemotaxis measurements in the same concentration and gradient range. The results here and in [4, 11] confirmed that the external noise dominates for both chemoattractants in low concentration range. This is in contrast with the conclusion in [33]. SI possibly caused by using single-pulse temporal gradients, as opposed to defined static gradients used here and in [4].

**ACKNOWLEDGMENTS**

We thank Petra Fey (Dicty Stock Center, Northwestern University) for AX4 cell lines, advice and suggestions about media preparation, Richard Kessin (Columbia University), Jeffrey Hadwiger (Oklahoma State University), David Knecht (University of Connecticut) for valuable advice relative to uniform receptor distributions for both FA [34] and cAMP receptors [35, 36]. Unlike in the cAMP case [4], here the non-circularity of cell shapes is not an issue since the cells are circular when sensing FA. However, there is still a possibility of a more complicated mechanism if FA receptors also transport FA into the cell [9], serving as a different communication channel, or if a FA transporter is a separate protein, as a separate communication channel. The possibility that FA and cAMP receptors share the majority of the internal signaling pathway [32] implies equal FA and cAMP responses, if rescaled by their respective parameters $K_d$ and $N$. This remains to be investigated with more cAMP and FA chemotaxis measurements in the same concentration and gradient range. The results here and in [4, 11] confirmed that the external noise dominates for both chemoattractants in low concentration range. This is in contrast with the conclusion in [33]. SI possibly caused by using single-pulse temporal gradients, as opposed to defined static gradients used here and in [4].
This work was supported by National Institutes of Health grant (P01 GM078586).


[22] We appreciate the comments of anonymous referees and Eric Siggia on this matter.


[24] We thank an anonymous referee for this suggestion.


METHODS

Information measures

Shannon’s information theory frees the data analysis from being tied to any particular model (as an example of successful applications see e.g. [1] and [2]) – and in this case, from any particular details of signal transduction pathways, but still provides quantifiable relationships between inputs and outputs. The relevant quantities in information theory are defined as follows [3]. The information entropy of a random variable $X$, is measured in bits defined as $H(X) = - \int p(x) \log_2 p(x) dx$ (a definite integral defined over the entire range where $X$ is defined). It is a measure of “sharpness” of probability distribution $p(x)$; a perfectly sharp probability distribution has entropy zero, whereas a perfectly flat, uniform distribution gives the highest possible value for entropy $H(X)$. An alternative interpretation of information entropy is the number of bits or the amount of information required to describe the random variable $X$. Sharp probability distributions require fewer bits for their full description than flat probability distributions. Intuitively, for the former only a few values near the peak can be sufficient to describe most of the outcomes of $X$, while for the latter we need more information to achieve the same. For conditional probability distributions, the conditional entropy is measured in bits defined as $H(X|Y) = - \int dy p(y) \int dx p(x|y) \log_2 p(x|y)$.

FIG. 1. The current paradigm for eukaryotic chemotaxis and the model assumptions. (a) Bacteria secrete folic acid (FA), which then binds to Dictyostelium FA receptors. Dictyostelium measures spatial distribution of occupied folic acid receptors and these binding events trigger a cascade of intracellular events eventually leading to cell movement. (b) Markov chain model assumption used in our work: the receptor occupancy $\theta_{rec}$ depends on the gradient $\theta_{grad}$, and the cell response $\theta_{res}$ conditionally depends on the gradient. (c) The external ($I_{ext}$) and total mutual information ($I_{tot}$) compared in this gradient. As detailed in Methods, $I_{ext}$ measures the information gained about the gradient, given the calculated spatial distribution of bound receptors, while $I_{tot}$ measures the information gained given the distribution of cell responses. Assuming the Markov chain relationship in part b), the data processing inequality states $I_{tot} \leq I_{ext}$.

Application to gradient sensing

In this case, the sensing process essential to eukaryotic chemotaxis is depicted in Fig. 1a. Here we consider three random but conditionally dependent variables, the gradient direction $\theta_{grad}$, the receptor occupancy $\theta_{rec}$ and the cell response directions $\theta_{res}$. These variables are assumed to form a Markov chain (see Fig 1b), where the cell response is conditionally dependent on the distribution of occupied receptors; i.e. given the distribution of occupied receptors, the cell response is completely independent of the original direction of the gradient that caused this particular receptor occupancy. Due to noise, the same receptor occupancy distribution can occur for gradients pointing in different directions. Without any prior knowledge we assume the gradient is equally likely to be pointing in any direction. We will see how much information we can obtain about the gradient by either observing the cell response and by calculating the distribution of receptor occupancy, and then comparing the two gains. The mutual information $I_{tot}(\theta_{grad}, \theta_{res}) = H(\theta_{grad}) - H(\theta_{grad}|\theta_{res})$ quantifies the total amount of information cells gained about the gradient (or by how much the entropy of $\theta_{grad}$ is reduced); this is determined by observing their response (see Fig 1c). Therefore, $I_{tot}$ is the gain in information that includes all possible noise sources in the FA signal transduction.
pathway.

In addition, the (external) mutual information (see Fig.1) between the gradient direction and receptor occupancy $I_{ext}(\theta_{grad}, \theta_{rec}) = H(\theta_{grad}) - H(\theta_{grad}|\theta_{rec})$ tells us the information gained about the gradient by knowing the distribution of receptors occupied with FA. Authors in [4] formulated a theory for computing this quantity and gave an analytical result applicable for shallow gradients. The assumptions behind this theory are: i) the steady state of the receptor-ligand binding process, ii) the first part of the Markov chain is not over a distance at which the ligand binding process is affected only by the local gradient, iii) cells of perfectly circular shapes and iv) uniform receptor distribution. While we have no direct way of confirming the plausible assumptions i) and ii) when sensing FA, Dictyostelium do have circular shapes and the distribution of FA receptors was previously measured as uniform [5]. This theory gives predictions for the external mutual information $I_{ext}$ using only two biochemical constants – the dissociation constant $K_d$ between FA and its receptor and the total receptor number per cell, $N$. Both have been measured previously and multiple receptor types/states have been discovered as is also the case for cAMP receptors (see main text for discussion). The dissociation constant and the total receptor number per cell, as well as the experimentally fixed FA concentration and its gradient in our devices are sufficient to predict the external mutual information $I_{ext}$. $I_{ext}$ provides the upper limit for the amount of information that can be acquired ($I_{tot}$), due to the data processing inequality: $I_{tot} \leq I_{ext}$ [7]. In other words, any kind of data processing can only destroy information. If the two quantities are roughly similar $I_{tot} \approx I_{ext}$, then the gain in information about $\theta_{grad}$ is about the same for the external mutual information $I_{ext}$. $I_{ext}$ is about the same for the majority of the noise in the entire process comes from receptor-ligand binding events.

**Cell growth and preparation**

Cells of the well characterized axenic strain, AX4 (provided by Dictyostelium Stock Center, Northwestern University), were grown in shaken culture suspension at 150 RPM in Formedium HL5 (Formedium, Hunstanton, UK) with glucose culture medium up to the concentration of about 0.5 – 3 × 10^6 cells. Development Buffer (DB; DictyBase recipe: 5 mM Na₂HPO₄, 5 mM KH₂PO₄, 1 mM CaCl₂, 2 mM MgCl₂; pH 6.5) was chosen as the medium for FA chemotaxis experiments because it is a well-defined medium and is an approximation of a physiological environment due to its low ionic strength [8]. A negative aspect of using DB is cell starvation and progression into development after 6+ hours (depending on cell density) and eventual loss of FA chemotactic sensitivity [9]. This was circumvented by performing the experiment before the starvation response occurs, as indicated by cell morphology – cells still had circular shapes. Since it was shown that the HL5 medium already contains about 0.12 mg/l of FA [10] (~0.3 μM), the medium was diluted by factor $\geq$30,000×, lowering the background FA concentration in the medium to at most 0.01 nM. This corresponds to about 1 molecule of FA per volume size of a Dictyostelium cell (100 μm^3). Depending on the cell concentration, 1-5 ml of cell suspension was taken from the shaken culture and DB was added for a total volume of 10 ml (dilution $\geq$2×). The cell suspension was then centrifuged for 40 seconds at 1000 RPM (200 g force), 9.8 ml of supernatant was removed, and 9.8 ml of DB was added to again have the final volume of 10 ml (dilution 50×); this was repeated once more (another dilution of 50×). 9.8 ml of supernatant was removed again and finally, 0.2 ml of 1μm diameter colloidal particles at concentration 10⁶ particles/ml (Polysciences, Inc.) in DB and 1-5 ml of DB was added, depending on the starting cell concentration (dilution $\geq$6×). The colloidal particles allowed us to monitor unintended convection that could ruin the static gradient. The entire procedure took about 20-30 minutes after which the cells were immediately loaded into the microfluidic device with an already established gradient.

**Microfluidics device design**

The microfluidic device was designed as an agarose gel containing 3 channels [11]: the static middle channel and two flowing side channels, that represent fixed boundary conditions, were separated by a layer of agarose gel and the gradient was formed by a layer of diffusion for FA to reach a steady state (see Fig.2 and Fig.3). Reservoirs were connected via Teflon tubing and the steady flow was supplied by a Harvard PHD 2000 syringe pump. The time to reach the steady state was checked by running a 2D diffusion simulation in COMSOL Multiphysics 3.5 (COMSOL, www.comsol.com) and analyzing the gradient in the middle of the channel (Fig.3). The microfluidic channel containing Dictyostelium cells, also contained 1μm-sized colloidal particles. These were used to monitor the flow rate in the static channel and the measured Peclet number $Le/D$ (dimensionless number characterizing the ratio of advective versus diffusive transport) was always below 0.3, where $L$ is the channel height (250 μm), $D$ the diffusion constant of folic acid 194 μm²/s [12] and $v$ the measured average drift velocity of colloidal particles (0.04 to 0.23 μm/s). After loading the cells, the gradient in the middle channel was temporarily lost, however, the time-scale of diffusive refilling of that channel from the bulk of the material above is estimated to be only $t \sim L^2/D \approx 5$ minutes, an insignificant duration.

**Device preparation**

The 3% agarose gel was formed as follows. 0.300g of agarose was mixed with 10 ml of DB. The agarose mixture was heated and kept near the boiling point in a microwave oven for 40 seconds total. Agarose was molded by pouring the heated mixture over an inverted PDMS master, which was itself molded from an original Teflon master produced by conventional milling. After about 2 minutes the agarose solidified, the holes were punched and the chamber was secured between a plexiglas manifold and a glass microscope slide. In this experiment 3% agarose serves as an environment permeable to small molecules, such as water and folic acid, but not permeable to Dictyostelium. Dictyostelium are
FIG. 2. A schematic of the microfluidic device used here.
migrating naturally attached on the glass surface, with 250 μm of static liquid (DB+FA gradient) on top and around them. The agarose gel was sealed well enough that the cells were unable to crawl underneath it.

Cell recording

For each run, at t=0 hours: the gradient formation was started. At t=3 hours: the cells were loaded in the device. Since we noticed that cells were not very mobile when first introduced into the device, we allowed them to adjust to the new environment for about 3.3 hours to establish a good degree of mobility. At t=6.3 hours recording started. At t=9.3 hours: the recording stopped. This time was chosen based on the fact that this is the time when one would first observe morphological changes associated with cell-to-cell cAMP signaling during the starvation response (e.g. elongated cells and formation of streams) when the cell density was significantly (10x) higher. Cell motion was recorded using bright field time-lapse optical microscopy, using an Olympus IX71 inverted microscope and a Home Science Tools MI-DC5000 5.0 Megapixel camera. Snapshots were taken every 30 seconds and cell trajectories were later analyzed on a computer. The list of concentrations used in both channels is shown in Table I.

Analysis of cell trajectories

We used ImageJ (http://imagej.nih.gov/ij/) with ParticleTracker Plugin [13] for automated cell detection and tracking. Particle tracks were analyzed in a custom-made MATLAB (The MathWorks, Natick, MA) code, where the following filtering was applied: the cells that could not be tracked consistently for more than 6 minutes (3% of the total recording time) were discarded and points on the screen that did not move at all were discarded as well; the latter corresponding to dead cells or other artifacts on the glass surface or CCD. Each experimental run was repeated 3 to 11 times, until about 300 to 700 cell trajectories were gathered. A sample of such trajectories is shown in Fig 4. The distribution of trajectories was very broad with lengths of 260 ± 220μm. Depending on the gradient, component of the velocity in gradient direction ranges from −0.15 μm/min to 0.51 μm/min.

Analysis of different trajectory time lengths

Here we experimentally check for the possibility that cells can integrate multiple gradient measurements over time scales longer than the pseudopod extension time (~ 30 s). We calculated the chemotactic index (CI) as we progressively moved the end point of the cell trajectory from the one at frame 2 (30 seconds) to the one at frame 400 (3.3 hours).

The 30 second time interval between subsequent frames was chosen since the cell displacements were typically about 3 μm, which was at the limit for measuring displacements in our experiments. If the cells were indeed integrating over more measurements as the time moved on, we would expect to see the CI

<table>
<thead>
<tr>
<th>combination</th>
<th>dc/dx (nM/μm)</th>
<th>c₀ (nM)</th>
<th>cₚ (nM)</th>
<th>cₚ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.20 × 10⁻⁴</td>
<td>5.0 × 10⁴</td>
<td>1.00 × 10⁵</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6.4 × 10⁰</td>
<td>1.0 × 10⁴</td>
<td>2.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3.2 × 10⁰</td>
<td>5.0 × 10³</td>
<td>1.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1.6 × 10⁰</td>
<td>2.5 × 10³</td>
<td>5.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>6.4 × 10⁻¹</td>
<td>1.0 × 10³</td>
<td>2.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>3.2 × 10⁻¹</td>
<td>5.0 × 10²</td>
<td>1.00 × 10³</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1.6 × 10⁻¹</td>
<td>2.5 × 10²</td>
<td>5.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>3.2 × 10⁻²</td>
<td>5.0 × 10¹</td>
<td>1.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>3.2 × 10⁻³</td>
<td>5.0 × 10⁰</td>
<td>1.00 × 10⁴</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>3.2 × 10⁻⁴</td>
<td>5.0 × 10⁻¹</td>
<td>1.00 × 10⁰</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>2.5 × 10³</td>
<td>2.50 × 10⁴</td>
<td>2.50 × 10³</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>1.0 × 10⁴</td>
<td>1.00 × 10⁵</td>
<td>1.00 × 10⁵</td>
</tr>
<tr>
<td>14</td>
<td>3.2 × 10⁻¹</td>
<td>5.0 × 10³</td>
<td>5.50 × 10⁴</td>
<td>4.50 × 10⁴</td>
</tr>
<tr>
<td>15</td>
<td>1.6 × 10⁰</td>
<td>7.5 × 10³</td>
<td>1.00 × 10⁴</td>
<td>5.00 × 10⁴</td>
</tr>
<tr>
<td>16</td>
<td>3.2 × 10⁰</td>
<td>1.5 × 10⁴</td>
<td>2.00 × 10⁵</td>
<td>1.00 × 10⁵</td>
</tr>
<tr>
<td>17</td>
<td>3.2 × 10⁰</td>
<td>5.0 × 10⁴</td>
<td>5.50 × 10⁵</td>
<td>4.50 × 10⁵</td>
</tr>
<tr>
<td>18</td>
<td>3.2 × 10⁻¹</td>
<td>5.0 × 10⁴</td>
<td>5.05 × 10⁵</td>
<td>4.95 × 10⁵</td>
</tr>
</tbody>
</table>

TABLE I. List of experimentally used concentrations in the two channels of a microfluidic device, c₁ and c₂ with calculated gradient dc/dx and the mean concentration c₀.
FIG. 4. Typical cell trajectories obtained from an experiment with $c_0 = 33 \, K_d$ and $dc/dx = 0.11 \, K_d/R$ ($K_d = 150 \, \text{nM}$). Different colors indicate different cell trajectories.

Fig.5 and $I_{ext,i} = I_{ext}(\langle c(x_i) \rangle)$ is the external mutual information for the average concentration in segment $i$. If $I_{ext}$ is averaged assuming a perfectly uniform cell distribution:

$$\langle I_{ext} \rangle = \frac{1}{c_{max} - c_{min}} \int_{c_{min}}^{c_{max}} I_{ext}(c_0)dc_0$$

(2)

the analytical result is:

$$\langle I_{ext} \rangle = \frac{N}{4 \ln 2 (c_{max} - c_{min})} \times$$

$$\left\{ \frac{1}{1 + c_{max}} - \frac{1}{1 + c_{min}} - \ln \left( \frac{1 + c_{min} c_{max}}{c_{min} (1 + c_{max})} \right) \right\}$$

which agrees to our estimate of $\langle I_{ext} \rangle$ to about 10% for our experiments.

FIG. 5. Average and the standard deviation of cell speeds for the experiments given in Fig.1b in the main text. $K_d = 150 \, \text{nM}$, $R = 5 \, \mu\text{m}$.

increase with time. The results for our peak experiments with the mean concentration of 2.5 $\mu\text{M}$ and the gradient of 1.6 $\text{nM}/\mu\text{m}$ are shown in Fig.7. Here we see that CI actually slightly decreases after $\sim 300$ s, but overall does not change significantly.

FIG. 6. Histogram showing the distribution of chemotactic index for our peak experiments with the mean concentration of 2.5 $\mu\text{M}$ and the gradient of 1.6 $\text{nM}/\mu\text{m}$.

FIG. 7. Chemotactic index (CI) as a function of the trajectory time length (or the maximum allowed integration time) for a single representative experimental run with the mean concentration of 2.5 $\mu\text{M}$ and the gradient of 1.6 $\text{nM}/\mu\text{m}$.

EFFECTS OF FOLIC ACID DEGRADATION

Here we explore the possibility that most of the FA is degraded by cells themselves, and they were effectively sensing a lower FA concentration, closer to $K_d$. FA can be degraded by an extracellular form of FA deaminase protein and we estimate the extent to which the FA concentration can be
fraction of cells

<table>
<thead>
<tr>
<th>0.00</th>
<th>0.02</th>
<th>0.04</th>
<th>0.06</th>
<th>0.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>-500</td>
<td>0</td>
<td>500</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIG. 8. The typical cell distribution as a function of the coordinate in the gradient direction, shown for the experimental run that gave the peak response with \( c_0 = 2500 \text{nM} \) and \( dc/dx = 1.6 \text{nM/μm} \) for \( M = 15 \) segments.

reduced by this process. Following up on the previous study of the level of deaminase secretion under the same conditions [14], we estimated the deaminase activity (defined as the amount of FA degraded per cell per unit time) for our experiment of total volume of 0.15 ml, about 50 cells in total and about 5 hours the cells spent in the chamber (corresponding to the middle of our run), the amount of FA that could possibly be degraded by that time is \( 5.25 \times 10^{-13} \text{mol} \). On the other hand, the total amount of FA in this entire volume, at 2.5 μM mean concentration is \( 3.75 \times 10^{-10} \text{mol} \), so the degradation by FA deaminase could account for less than 0.1% of the expected amount of FA. This calculation is summarized in the Table II. This conclusion was verified experimentally by changing the cell density by a factor of four (from 7 cells/mm² to 30 cells/mm²) for the gradient where we observed peak response and noticing that the same result in terms of chemotactic index (0.10 ± 0.02 at lower vs 0.09 ± 0.01 at higher density) and total mutual information (0.14 ± 0.02 bits vs 0.14 ± 0.01 bits) was observed. Thus, we conclude that degradation of FA by FA deaminase cannot account for the violation of the data processing inequality.

### Table II. Summary of the calculation for FA deaminase contribution to the observed results for the case of our best response at 2.5 μM mean concentration.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>35 × 10^{-6}</td>
<td>pmol/(cell min)</td>
</tr>
<tr>
<td>Total volume</td>
<td>0.15</td>
<td>ml</td>
</tr>
<tr>
<td>Time</td>
<td>300</td>
<td>min</td>
</tr>
<tr>
<td>Cell number</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>FA amount</td>
<td>3.75 × 10^{-10}</td>
<td>mol</td>
</tr>
<tr>
<td>FA amount degraded</td>
<td>5.25 × 10^{-13}</td>
<td>mol</td>
</tr>
<tr>
<td>FA percentage degraded</td>
<td>0.07</td>
<td>%</td>
</tr>
</tbody>
</table>

EFFECTS OF CELL POLARIZATION / BIAS

The total mutual information with bias is defined by:

\[
I_{\text{tot}}^{\text{bias}} = H_{\text{bias}}(\theta_{\text{res}}) - H_{\text{bias}}(\theta_{\text{res}}|\theta_{\text{grad}})
\]

with

\[
H_{\text{bias}}(\theta_{\text{res}}) = - \int p(\theta_{\text{res}}; K) \log_2 p(\theta_{\text{res}}; K) d\theta_{\text{res}}
\]

\[
I_{\text{bias}}^{\text{ext}} = I_{\text{ext}} - B(K)
\]

\[
I_{\text{bias}}^{\text{ext}} = I_{\text{ext}} - \int p(\rho) h(\rho; K_p) d\rho
\]

with:

\[
I_{\text{ext}} = \frac{1}{\ln 2} \left( \frac{\nu}{\sigma} \right)^2 - \int p(\rho) \log_2 J_0 \left( \frac{\rho \nu}{\sigma^2} \right) d\rho
\]

\[
p(\rho; \nu, \sigma) = \frac{\rho}{\sigma^2} \exp \left[ -\frac{\rho^2 + \nu^2}{2\sigma^2} \right] I_0 \left( \frac{\rho \nu}{\sigma^2} \right)
\]

\[
\nu(N, c_0, \nabla c) = \frac{N}{2} \frac{\nabla c}{c_0 + 1}
\]
The total mutual information calculated using Eq.3 (main text) depends on the choice of number of bins $m$.

While there is no “best” number of bins, here the total number of bins chosen was 14 which gave similar results for all combinations of gradients and mean concentrations since we had roughly the same number of cells in each case (typically around 500). First, as stated in the main text, it correlates well with the CI (comparing Fig.1b and 2a in the main text). Second, $I_{tot}$ reaches a plateau in this bin range and becomes lower when we use too few bins (below $\approx 10$) or higher but with much larger uncertainty if we use too many bins (roughly 30 or more); see Fig[10] The plateau corresponds to the middle ground here where $I_{tot}$ does not change much if the bin number changes a little around the chosen value. Finally, this choice of 14 bins gave approximately the same results as the Kernel Density Estimate (Fig.9) used for data smoothing [15].