Model for cooperative control of positional information in *Drosophila* by *bicoid* and maternal *hunchback*.

John Reinitz*, Eric Mjolsness†, and David H. Sharp‡

* Center for Medical Informatics

Yale Medical School, 333 Cedar St., New Haven, CT 06510 USA

†Department of Computer Science, Yale University

P.O. Box 2158 Yale Station, New Haven CT 06520 USA

‡Theoretical Division, Los Alamos National Laboratory

Los Alamos, NM 87545 USA

Abstract

The blastoderm of the fruit fly *Drosophila melanogaster* is unusually well suited for analysis of fundamental questions in animal development. One such question is how genes specify the positional information which determines the developmental pathways (fate) of cells at appropriate spatial locations. In this paper we propose a dynamical model of gene regulation which explicitly describes how positional information is used in the blastoderm. The model is applied to analyze important experimental findings on the dependence of cell fate on the concentration of the *bicoid* morphogen. The model shows that positional information in the presumptive middle body is cooperatively determined by maternal products of the *bicoid* and *hunchback* genes.

Cells in an embryo organize themselves into appropriate structures at correct locations by interpreting positional information [1] encoded in chemical signals. The purpose of this paper is to propose a model of gene regulation which provides an explicit dynamical formulation of the notion of "interpretation of positional information". The model is developed with reference to the blastoderm of the fruit fly *Drosophila melanogaster*. In this system, positional information is supplied by maternal genes and interpreted by zygotic genes. On the anterior-posterior (A-P) axis, the interpretation of positional information is manifested by the expression of finer and finer spatial domains of zygotic segmentation gene expression [2] under the control of three maternal organizing centers [3, 4]. The two of these that operate in the middle body region have been shown to act by means of morphogen gradients: *bicoid* (*bcd*) in the case of the anterior system [5, 6, 7], and maternal *hunchback* (hb^{mat}) in the abdominal system [8, 9, 10, 11, 12].

If positional information were specified by a single chemical gradient of morphogen, one would expect that a given morphological marker of cell fate would always be found at a given concentration of morphogen. This is not the case in *Drosophila*. An important series of quantitative measurements ([7], Fig. 6) have shown that the same cell fate can be elicited by different concentrations of *bcd*. In this paper we use our model to show that A-P positional information is encoded jointly by the *bcd* and hb^{mat} gradients. The model correctly predicts the anomalous displacements in expression pattern observed in [7], and makes predictions about the expression of gap genes under various doses of *bcd* in embryos lacking hb^{mat}.

Formulation of the Model

Our model incorporates certain fundamental experimental observations on the development of *Drosophila melanogaster*. Until the onset of gastrulation, the *Drosophila* embryo has two unusual features. One is the absence of cells: the syncytial character of the blastoderm permits spatial interactions to be treated by the diffusive exchange of gene products, and explicit cell-cell interactions are not present [13]. Second is the observation that although mutation of a segmentation gene alters the expression patterns of other segmentation genes [14, 15, 16, 17, 18, 19, 20], it does not produce an alteration in the morphology of the embryo until some time after gastrulation [21, 22, 23, 24]. These observations show that segmentation genes are dynamically coupled to each other, but are isolated from other developmental processes until after gastrulation. Between fertilization and gastrulation 13 synchronous nuclear divisions occur, but no zygotic genes are expressed until after the tenth division (cleavage cycle 11) [25]. In this paper we focus on a particular class of segmentation genes known as gap genes [26]. These genes are the major targets of maternal morphogens and themselves regulate a cascade of pair-rule and segment polarity genes which generate a segmental pattern that is visible at the molecular level by the time of gastrulation [2] and at the morphological level by the time of germ band elongation [13].

We make two hypotheses about the way that segmentation genes interact. Hypothesis (a) is that the state of the blastoderm can be described by the concentrations of gene products, and (b) is that the synthesis rate of a given gene is a monotonic, saturating ("sigmoidal") function of the regulating gene products.

Hypothesis (a) is based on the idea that precise binding states of regulator proteins are not required for a description of the regulative state of a gene. Although the regulative state is in some sense a consequence of ligand binding, these binding configurations are highly redundant and thus do not provide a useful way of specifying the regulative state of the gene. A consequence of this picture is that the action of one gene on another can be described simply in terms of its effect on concentrations, and hence by a single real number.

Hypothesis (b) is supported by some *in vivo* data [27], but decisive experimental information on the form of this function must await a faithful *in vitro* assay for RNA polymerase II regulation that can be performed with purified transcription factors.

We formulate these basic ideas as a mathematical model in the following way. Concentrations of regulatory molecules change in response to existing concentrations of regulators, exchange of regulatory molecules between nuclei (by diffusion), and decay. These effects are described in the model by a system of coupled nonlinear differential equations. The system of differential equations holds at almost all times, but when mitosis occurs the state of the system is considered to change discontinuously. Such discrete events are governed in our model by a rule which specifies how a given nucleus is replaced by its progeny. This rule is part of a grammar in the sense of Lindenmayer [28, 29]. We have designed our model to take into account the fact that gene expression takes place only during interphase, and that each nuclear division is symmetric and lasts about four minutes. Because the timing of mitoses in the blastoderm is under maternal control, we impose a fixed schedule of mitoses on the model [30].

We consider the four gap genes *Kruppel* (*Kr*), *knirps* (*kni*), *giant* (*gt*), and *hunchback* (*hb*) under the control of *bcd* and focus on that portion of the blastoderm which gives rise to the segmented germ band. In this region, the level of expression of gap genes is approximately a function only of the position along the A-P axis. Moreover the genes affecting the A-P and dorsal-ventral (D-V) axes are uncoupled, because mutations in D-V genes do not affect the expression of A-P genes and *vice versa*. These observations suggest that the system can be well represented by a linear array of nuclei running along the A-P axis. The position of a nucleus is indexed by i. During interphase the dynamical equation for v_i^a, the concentration of the ath gene product in nucleus i, is given by [29, 31]

$$\frac{dv_i^a}{dt} = R_a g_a(\sum_{b=1}^{N} T^{ab} v_i^b + m^a v_i^{bcd} + h^a) + D^a(n)\Big[(v_{i-1}^a - v_i^a) + (v_{i+1}^a - v_i^a)\Big] - \lambda_a v_i^a, \qquad (1)$$

where N is the number of zygotic genes.

In (1), T^{ab} is the matrix of genetic regulatory coefficients; the elements of T^{ab} characterize the regulatory effect of gene b on gene a. This matrix does not depend on i, a reflection of the fundamental fact that the cell nuclei of a multicellular organism contain identical genetic material. The *bcd* input is given by $m^a v_i^{bcd}$, where v_i^{bcd} is the concentration of *bcd* protein in nucleus i and m^a is the regulatory coefficient of *bcd* acting on zygotic gene a. g_a is a monotonic, saturating function of the form $g_a(u^a) = (1/2)[(u^a/\sqrt{(u^a)^2+1}) + 1]$, where $u^a = \sum_{b=1}^{N} T^{ab} v_i^b + m^a v_i^{bcd} + h^a$. R_a is the maximum rate of synthesis from gene a, h^a is a threshold that controls the constitutive activity of gene a. The diffusion parameter $D^a(n)$ depends on the number n of cell divisions that have taken place, and varies inversely with the square of the distance between nuclei. In this paper, we take $D^a(n)$ to be the same for each species of protein. We assume that the distance between adjacent nuclei is halved after a nuclear division. λ_a is the decay rate of the product of gene a.

hb^{mat} protein concentration is incorporated into the model as an initial value (that is at cleavage cycle 11) for the concentration of *hb* product. This is a mathematical expression of the fact that the observed concentration of *hb* protein consists of both maternal and zygotic components. Initial values of the other, purely zygotic, gene products are taken to be zero at cleavage cycle 11.

How the Model is Used

The first step in using the model is to employ experimental data on levels of gene expression to determine the parameters in equation (1). Note that this is an inverse procedure; it is not that we are starting with the values of T^{ab} and other parameters to determine $v^a(t)$, but rather we are using the experimentally observed values of $v^a(t)$ to define the parameters. This procedure is carried out in two steps. One must first analyze the experimental data on levels of gene expression to obtain quantitative values for the protein concentrations as a function of time, and secondly one must use this data to fit the parameters by a least squares procedure which is solved by the method of simulated annealing [32, 33].

Experimental data derived from single and double labeling studies on wild type embryos using fluorescence tagged antibodies were used to construct the dataset shown in Fig. 1 [34]. The figure displays concentrations of the products of *hb*, *Kr*, *kni*, and *gt* in early, middle, and late cleavage cycle 14, as well as the concentration profile of *bcd* [35], which is not a function of time. The concentration of hb^{mat}, together with the other initial conditions of the model, are shown in Fig. 2.

To use the data in Fig. 1 to determine the parameters in equation (1), we form the sum of squared differences (the "cost function") between the observed protein concentrations and those given by the model. The sum extends over each protein, nucleus, and time for which data exists. This quantity is then minimized by adjusting the values of the parameters. This defines a least squares problem which is solved using Lam's adaptive simulated annealing algorithm [36, 37, 38].

This procedure led to the results shown in Fig. 2. Comparison of Fig. 1 and Fig. 2 shows that the relative spacing and overlaps of expression domains is correct for the *Kr* and *kni* domains, for the anterior and posterior *hb* domains, and for the posterior portion of the anterior domain of *gt* and the anterior border of the posterior *gt* domain. These points of agreement are the basis for our claim that the model is correct in most of the region of the blastoderm under consideration. In the anterior and posterior end regions, the model gives results that are variable over different simulated annealing runs and partially incorrect. This is not unexpected, because genes which are known to have an important regulatory effect in these regions have not yet been included in the model [39, 40, 41, 42]. For example, in cleavage cycle 12, there was an unbiologically large transient in *kni* expression in the posterior region. This spurious early *kni* transient activates the posterior *hb* domain and represses the posterior *gt* domain, regulatory actions that are characteristic of the gap gene *tailless* [15], which is expressed in this region [43].

The importance of including all relevant genes was further illustrated by performing fits in the absence of *bcd* (Fig. 3). As can be seen by comparing Figs. 2 and 3, the fit without *bcd* contains major qualitative errors, notably the loss of the anterior border of the anterior *gt* domain as well as spurious domains of *Kr*, *kni*, and *hb* expression posteriorly. This result supports the model, because *bcd* activity is required to set gap domains in the anterior half of this region [5]. This illustrates the important point that the model does not give a good result when it should not: A good fit is obtained only when the proper gene products are included.

The method used to obtain the data (Fig. 1) gives good information about the relative overlap of expression domains. However, potentially serious uncertainties in the estimated values of protein concentrations are present for several reasons: The levels of expression were derived by visual inspection, the relation between observed intensity and concentration may not be linear, and the intensity-concentration relationship may be different for each gene. We emphasize that there is reason to believe that many key biological properties of this system are closely related to the overlap of expression domains, and are relatively insensitive to uncertainties in absolute concentrations. For example, the ordered set of gap gene expression domain overlaps appears experimentally to specify the location of pair-rule stripes, and theoretical analysis [44] indicates that this set of overlaps lay out a periodic set of positional cues for segmentation. Absolute concentrations of gap gene products cannot be critical, since reducing the concentration of a gap gene product by a factor of two in a heterozygote merely changes the relative spacing of the overlaps [45], without an effect on viability [26]. Elimination of an overlap (as in a null gap mutant) has lethal results and eliminates certain fates from the fate map. Alterations in the positions of overlaps as is observed in *bcd* dosage studies may change the fates of many blastoderm nuclei, but no fates are eliminated and the embryo survives [7].

Analysis of *bcd* Dosage Studies

Dosage experiments which change the number of copies of the *bcd* gene provide a way of determining the dependence of positional information on the chemical gradient of *bcd* product. This is accomplished by observing the response of suitable markers to changes in the concentration of Bicoid protein. Published photometric results [7] indicate that the gradient remains close to exponential for 1, 2, or 3 doses of *bcd*, and that the absolute scale of the gradient is roughly a linear multiple of the copy number in the part of the embryo under consideration here. In this study, the parameters of the model are fit on wild type data only, with the normal dosage of two copies of *bcd*. We can model the effect of alterations in the *bcd* dose without redetermining the model parameters simply by multiplying the *bcd* gradient used by an appropriate scale factor [46].

If A-P positional information were autonomously specified by the concentration of *bcd*, one would expect that a given landmark used to assay positional information would always be found at the same concentration of *bcd* product. The authors of [7] used the cephalic furrow and the first *even-skipped* (*eve*) stripe as such markers. Although these markers changed position with changes in the *bcd* dose, they did not remain at the same *bcd* concentrations (Fig. 4, curve d), indicating that *bcd* does not autonomously specify positional information.

Our model correctly predicts this behavior. Because neither pair-rule genes or gastrulation was included in the model, we have used the anterior edge of the *Kr* domain at the end of cleavage cycle 14 as a positional information marker. This is a suitable marker because it is invariant with respect to the fate map for the range of genotypes considered. This is true for *Kr* at the end of cycle 14 because its position is invariant with respect to the *eve* pair-rule stripes: The anterior margin of the *Kr* domain demarcates the posterior boundary of the second *eve* stripe [47]. Fig. 4 shows that the model predicts the same relationship between *bcd* concentration and positional information as described by the slightly more anterior markers used in [7].

To determine the contribution of hb^{mat} to the specification of positional information, we also fit the model to a dataset that did not include hb^{mat}, using data otherwise identical to that used for the fit shown in Fig. 2. As shown in Fig. 4 the results demonstrate that in the absence of hb^{mat} positional information is determined solely by the *bcd* concentration. This, together with the dose response in wild type, implies that in wild type embryos positional information in the presumptive thoracic and anterior abdominal region of the blastoderm is determined cooperatively by both *bcd* and hb^{mat}.

A possible objection to the preceding analysis is that it should have been based on expression patterns in embryos without hb^{mat}; nevertheless we believe our conclusions to be valid for the following reasons. Embryos lacking hb^{mat} survive to become viable larvae [21], and experimental data [11, 12] supports the idea that gap gene expression in such embryos will contain the same sequence of gap expression domains and their overlaps as wild type. Because our conclusions about positional information in the absence of hb^{mat} are drawn purely from the changes in the positioning of the domains of the four gap genes considered, rather than their absolute positions, we believe that our analysis is not sensitive to distortions in the gap pattern introduced by the absence of hb^{mat}.

We note that our model leads to the experimental prediction that if the *bcd* dosage is varied between 1 and 3 in embryos lacking hb^{mat}, the same gap gene expression domains will occur at the same concentration of *bcd*. We emphasize that this prediction depends on the assumption that the *torso* system, which has not yet been included in our model, does not play an important role in this phenomenon. This assumption is based on the fact that the landmark used in Fig. 4 is at the border of the regions organized by the anterior and abdominal maternal systems, and hence is distant from known sites of *torso* action.

Conclusions

We have presented results which support the idea that a model based on a simple picture of gene regulation can successfully describe important features of gene expression in the blastoderm of *Drosophila*. We have applied this model to elucidate an important aspect of the control of positional information, and to predict the results of further experiments. The inclusion of more genes in the model will bring many more phenomena within its framework. We have identified uncertainties in the determination of the model parameters resulting from the semiquantitative nature of data obtained by visual inspection of antibody stains. These can be reduced substantially if data are collected using charge coupled device (CCD) or confocal techniques. This more quantitative data will also permit modeling of many experiments on single and double null mutants.

References

[1] Wolpert, L., *Journal of Theoretical Biology* **25**, 1–47 (1969).

[2] Ingham, P. W., *Nature* **335**, 25–34 (1988).

[3] Nusslein-Volhard, C., Frohnhofer, H. G., and Lehmann, R., *Science* **238**, 1675–1687 (1987).

[4] Nusslein-Volhardt, C., *Development Supplement* **1**, 1–10 (1991).

[5] Frohnhofer, H. G. and Nusslein-Volhardt, C., *Nature* **324**, 120–125 (1986).

[6] Driever, W. and Nusslein-Volhard, C., *Cell* **54**, 83–93 (1988).

[7] Driever, W. and Nusslein-Volhard, C., *Cell* **54**, 95–104 (1988).

[8] Hulskamp, M., Schroder, C., Pfeifle, C., Jackle, H., and Tautz, D., *Nature* **338**, 629–632 (1989).

[9] Irish, V., Lehmann, R., and Akam, M., *Nature* **338**, 646–648 (1989).

[10] Struhl, G., *Nature* **338**, 741–744 (1989).

[11] Hulskamp, M., Pfeifle, C., and Tautz, D., *Nature* **346**, 577–580 (1990).

[12] Struhl, G., Johnston, P., and Lawrence, P. A., *Cell* **69**, 237–249 (1992).

[13] Campos-Ortega, J. A. and Hartenstein, V., *The Embryonic Development of* Drosophila melanogaster, Springer-Verlag, Berlin, 1985.

[14] Jackle, H., Tautz, D., Schuh, R., Seifert, E., and Lehmann, R., *Nature* **324**, 668–670 (1986).

[15] Reinitz, J. and Levine, M., *Developmental Biology* **140**, 57–72 (1990).

[16] Mohler, J., Eldon, E. D., and Pirotta, V., *The EMBO Journal* **8**, 1539–1548 (1989).

[17] Pankratz, M. J., Hoch, M., Seifert, E., and Jackle, H., *Nature* **341**, 337–340 (1989).

[18] Kraut, R. and Levine, M., *Development* **111**, 601–609 (1991).

[19] Kraut, R. and Levine, M., *Development* **111**, 611–621 (1991).

[20] Eldon, E. D. and Pirrotta, V., *Development* **111**, 367–378 (1991).

[21] Lehmann, R. and Nusslein-Volhard, C., *Developmental Biology* **119**, 402–417 (1987).

[22] Wieschaus, E., Nusslein-Volhard, C., and Jurgens, G., *Roux's Archives of Developmental Biology* **193**, 296–307 (1984).

[23] Petschek, J. P., Perrimon, N., and Mahowald, A. P., *Developmental Biology* **119**, 175–189 (1987).

[24] Strecker, T. R., Merriam, J. R., and Lengyel, J. A., *Development* **102**, 721–734 (1988).

[25] Edgar, B. and Schubiger, G., *Cell* **44**, 871–877 (1986).

[26] Nusslein-Volhard, C. and Wieschaus, E., *Nature* **287**, 795–801 (1980).

[27] Han, K., Levine, M. S., and Manley, J. L., *Cell* **56**, 573–583 (1989).

[28] Lindenmayer, A., *Journal of Theoretical Biology* **18**, 280–315 (1968).

[29] Mjolsness, E., Sharp, D. H., and Reinitz, J., *Journal of Theoretical Biology* **152**, 429–453 (1991).

[30] Foe, V. A. and Alberts, B. M., *Journal of Cell Science* **61**, 31–70 (1983).

[31] Reinitz, J., Mjolsness, E., and Sharp, D. H., A Connectionist Model of the *Drosophila* Blastoderm, in Mittenthal, J. E. and Baskin, A. B., editors, *The Principles of Organization in Organisms*, Addison-Wesley, 1992.

[32] Metropolis, N., Rosenbluth, A., Rosenbluth, M. N., Teller, A., and Teller, E., *Journal of Chemical Physics* **21**, 1087–1092 (1953).

[33] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., *Science* **220**, 671–680 (1983).

[34] Data shown in Fig. 1 B, C, and D were obtained from single and double labeling studies of wild type (Oregon R) embryos using fluorescence tagged antibodies, with fixation and staining [48] and photomicroscopy [47] as described. The *Kr* antiserum was that used [47], the *kni* antiserum was obtained from Dr. K. Howard, and guinea pig anti-*gt* and mouse anti-*hb* sera were those described [18]. These studies result in color photomicrographs, each containing expression patterns of at most two genes at a particular stage of development. In order to be used as input to the model, these data must be analyzed to obtain quantitative values for protein concentrations and relative positions of expression domains. To obtain concentrations as a function of time, embryos were classified by developmental stage. Visual estimates of expression intensity along a line of nuclei running along the A-P axis close to the lateral midline were made from each embryo.

The alignment problem was dealt with by comparing data from different embryos which have a common expression domain. Note that the late cycle 14 dataset is truncated at the anterior end owing to the fact that we did not have adequate expression data in this region.

[35] *bcd* concentration data was taken from photometric observations in the literature [6]. The curve shown is proportional to $\exp(-.05x)$, where x is the nucleus number.

[36] Lam, J. and Delosme, J.-M., An efficient simulated annealing schedule: derivation, Technical Report 8816, Yale Electrical Engineering Department, New Haven, CT, 1988.

[37] Lam, J. and Delosme, J.-M., An efficient simulated annealing schedule: implementation and evaluation, Technical Report 8817, Yale Electrical Engineering Department, New Haven, CT, 1988.

[38] A general description of simulated annealing is given in Kirkpatrick *et al.* [33]. In the present application, evaluation of the cost function requires integration of equation (1) for each trial set of T^{ab}, R^a, etc. to obtain $v^a_{model}(t)$. Eq (1) was integrated with 1.33 minute time steps using the Euler method [49]. Experiments with finer timesteps confirmed the accuracy of the solution. Annealing moves for each parameter were selected from an exponential distribution and given random sign. The mean of this distribution was separately altered for each parameter so as to ensure an effective sampling of search space. The search space for parameters is limited as follows. We restrict the synthesis rate $Rg(u)$ to be no more than 99.9% of its maximum value. We have used two strategies for annealing on R and λ. In the first, λ was fixed at a value corresponding to a protein half life of 20 minutes, and R was allowed to range from 0.25 min $^{-1}$ to 2.5 min $^{-1}$. In

the second, λ assumed values corresponding to half lives ranging from 8 to 24 minutes, and R ranged from 0.25 min $^{-1}$ to 0.75 min $^{-1}$. The latter search space gave scores about 10% better than the former. In each of the two search spaces, multiple annealing runs were done and those with scores within 15% of the lowest score in each set were examined in detail. Lam's parameter which determines the overall cooling rate [37] was set to 10^{-4}. Slower cooling did not improve the results. We have tested the accuracy of this solution algorithm in the following way. Starting with an arbitrarily chosen set of parameters T^{ab}, R^a etc. one solves equation (1) to obtain a set of concentrations $v_i^a(t)$. These concentrations are regarded as the "experimental" input data. The simulated annealing procedure is then carried out to determine the set of parameters determined by this input data. If these parameters agree with those initially chosen, the solution is accurate. This test was carried out for a number of systems; for a two gene system it was found that the initial and final values of the parameters agreed to within 1%, and for the 4 gene system considered here most parameters agreed to within 10 %, although in a few cases only 50 % agreement was obtained.

[39] Dalton, D., Chadwick, R., and McGinnis, W., *Genes and Development* **3**, 1940–1956 (1989).

[40] Weigel, D., Jurgens, G., Klingler, M., and Jackle, H., *Science* **248**, 495–498 (1990).

[41] Finkelstein, R. and Perrimon, N., *Nature* **346**, 485–488 (1990).

[42] Cohen, S. M. and Jurgens, G., *Nature* **346**, 482–485 (1990).

[43] Pignoni, F., Baldarelli, R. M., Steingrimsson, E., Diaz, R. J., Patapoutian, A., Merriam, J. R., and Lengyel, J. A., *Cell* **62**, 151–163 (1990).

[44] Kauffman, S. and Goodwin, B., *Journal of Theoretical Biology* **144**, 321–345 (1990).

[45] Warrior, R. and Levine, M., *Development* **110**, 759–767 (1990).

[46] Throughout this analysis changes in the dosage of *bcd* were represented numerically as multiples of the wild type *bcd* gradient [35]. The solid curve in Fig. 4 was obtained from the formula $y_0 = A(n/2)\exp(-ax_0(n))$, where $y_0 = 2.35$ is the fixed *bcd* protein concentration at the posterior margin of the *Kr* domain for $n = 2$, n is the dosage of *bcd* (interpreted as a continuous variable), and $x_0(n)$ the location of the anterior edge of the *Kr* domain as a function of dosage. $A = 9.51$ is the relative concentration of *bcd* at nucleus 0 for $n = 2$, $a = .05$ the decay constant of the *bcd* gradient (Fig. 1 A). In curves a,b,c, and e of Fig. 4, the anterior edge of the *Kr* domain was taken to be the position at half maximal amplitude. Occasionally a fit lost or gained expression domains at 1 or 3 doses; these were not used to make dose response curves. Curve b was obtained from the model shown in Fig. 2. The curve with the hollow squares was obtained by fitting to data identical to that shown in Fig. 1, except that the initial *hb* concentration was set equal to zero. The fit had an rms deviation of 1.24 intensity units and the parameters were $T^{Kr\leftarrow Kr} = 1.76$, $T^{Kr\leftarrow hb} = -1.00$, $T^{Kr\leftarrow gt} = 1.31$, $T^{Kr\leftarrow kni} = -0.585$, $T^{hb\leftarrow Kr} = -0.204$, $T^{hb\leftarrow hb} = 0.922$, $T^{hb\leftarrow gt} = -0.249$, $T^{hb\leftarrow kni} = -0.285$, $T^{gt\leftarrow Kr} = -3.84$, $T^{gt\leftarrow hb} = -0.136$, $T^{gt\leftarrow gt} = -0.045$, $T^{gt\leftarrow kni} = -0.825$, $T^{kni\leftarrow Kr} = -0.452$, $T^{kni\leftarrow hb} = -0.416$, $T^{kni\leftarrow gt} = -2.149$, $T^{kni\leftarrow kni} = 0.299$, $h^{Kr} = -7.10$, $h^{hb} = -1.64$, $h^{gt} = 1.29$, $h^{kni} = 3.04$, $t^{Kr}_{1/2} = 21.1$ min., $t^{hb}_{1/2} = 11.4$ min., $t^{gt}_{1/2} =$

12.0 min., $t_{1/2}^{kni} = 14.2$ min., $R^{Kr} = 0.416\ \mathrm{min}^{-1}$., $R^{hb} = 0.711\ \mathrm{min}^{-1}$., $R^{gt} = 0.706\ \mathrm{min}^{-1}$., $R^{kni} = 0.614\ \mathrm{min}^{-1}$., $m^{Kr} = 0.459$, $m^{hb} = 0.493$, $m^{gt} = 0.592$, $m^{kni} = 0.033$, $D(13) = 0.032\ \mathrm{min}^{-1}$..

[47] Stanojevic, D., Hoey, T., and Levine, M., *Nature* **341**, 331–335 (1989).

[48] Frasch, M. and Levine, M., *Genes and Development* **1**, 981–995 (1987).

[49] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1988.

[50] JR acknowledges support from National Institutes of Health grants LM 07056 and RR 07801, and from Sun Microsystems Inc.; EM from the US Air Force Office of Scientific Research grant 88-0240, and DHS from National Institutes of Health grant RR 07801 and from the US Department of Energy. JR thanks Michael Levine and members of his laboratory for stimulating discussions, and Daniel Greening for advice on simulated annealing. We thank K. Harding and G. Zina for valuable comments on the manuscript.

Figure Legends

FIG. 1 Dataset used to fit the model. A. Gradient of maternal *bcd* protein in cycle 14 [35]. The *bcd* concentration was kept constant in time. B, C, and D. Zygotic expression data from early, middle, and late cleavage cycle 14 [34]. Each shows a line of nuclei along the lateral midline, with anterior to the left. The same region of 64 nuclei is shown in each graph. Concentration is plotted vertically: currently available data only permit the determination of relative concentrations. Domains are labeled in C; + signs denote *hb*, the x signs *kni*, the diamonds *Kr*, and the open squares *gt*.

FIG. 2 Behavior of the model. The figure shows the behavior of the model from cleavage cycle 11 to the onset of gastrulation. Symbols and axes are as in Fig. 1, except for altered horizontal scales in A through C. A shows initial conditions in cycle 11. Note that all gene products are at zero concentration, except that of *hb*; *hb* product in cycle 11 was expressed from the maternal genome. The concentration profile of hb^{mat} was taken directly from data; no other expression was seen at this time. B, C, D, E, and F show the behavior of the model at cleavage cycles 12, 13, early, middle, and late cycle 14 respectively. This fit [38] was the best obtained from the data in Fig. 1, with a root mean square (rms) deviation of 1.24 intensity units. The fit gave the following parameters: $T^{Kr\leftarrow Kr} = 0.997$, $T^{Kr\leftarrow hb} = -0.826$, $T^{Kr\leftarrow gt} = -1.71$, $T^{Kr\leftarrow kni} = -1.93$, $T^{hb\leftarrow Kr} = 0.442$, $T^{hb\leftarrow hb} = 0.998$, $T^{hb\leftarrow gt} = 1.06$, $T^{hb\leftarrow kni} = 0.398$, $T^{gt\leftarrow Kr} = -3.58$, $T^{gt\leftarrow hb} = -0.298$, $T^{gt\leftarrow gt} = 0.190$, $T^{gt\leftarrow kni} = 0.284$, $T^{kni\leftarrow Kr} = -0.081$, $T^{kni\leftarrow hb} = -1.05$, $T^{kni\leftarrow gt} = -1.51$, $T^{kni\leftarrow kni} = 0.258$, $h^{Kr} = 0.791$, $h^{hb} = -6.72$, $h^{gt} = -1.16$, $h^{kni} = 2.66$, $t_{1/2}^{Kr} = 24.0$ min., $t_{1/2}^{hb} = 16.1$ min., $t_{1/2}^{gt} = 11.0$ min., $t_{1/2}^{kni} = 10.5$ min., $R^{Kr} = 0.298$ min^{-1}., $R^{hb} = 0.511$ min^{-1}., $R^{gt} = 0.700$ min^{-1}., $R^{kni} = 0.749$ min^{-1}., $m^{Kr} = 1.59$, $m^{hb} = 0.029$, $m^{gt} =$

1.19, $m^{kni} = -1.55$, $D(13) = 0.087\ \mathrm{min}^{-1}$..

FIG. 3 Requirement for *bicoid.* This figure shows the results of fitting the model to the data in Fig. 1 with the *bcd* gradient set to zero. Late cleavage cycle 14 is shown; symbols and axes are as in Fig. 1, and expression domains are labeled. The results are qualitatively incorrect in several respects; compare Fig. 1 D and 2 F and in this figure note the ectopic expression of *Kr*, *kni*, and *hb* in the posterior region. Computational methodology is as described [38]. This fit was the best obtained without *bcd.* It had an rms deviation of 1.54 intensity units, parameters not shown.

FIG. 4 The *bicoid* dose response. This figure [46] shows that the model correctly predicts the observed *bcd* dose response curve and also gives the predicted dose response curve in the absence of hb^{mat}. Gene dosage is plotted on the vertical axis in terms of the number of copies of the wild type *bcd* gene that are present. The horizontal axis is in the same units as Figs. 1 through 3; note that one nucleus is approximately 1% of egg length. Five curves are shown. The solid curve a shows the location of a fixed concentration of bicoid protein at different doses of the *bcd* gene as given by an amplitude scaling of the exponential formula (Fig. 1) for the *bcd* concentration. This curve is to be compared to curve e (hollow squares) which shows the variation in position of the anterior margin of the *Kr* domain as a function of *bcd* dosage when hb^{mat} was not included in the model. There was no significant variation in this curve among all the fits examined. Curves b (diamonds) and c (+ symbols) show the variation in position of the anterior margin of the *Kr* domain as a function of *bcd* dosage when both *bcd* and hb^{mat} are included in our model (with parameters for curve b given in Fig. 2; parameters for curve c not shown). Curves b and c are to be compared to curve d (x symbols), which represents the published experimental data [7] on the dose response.

Curves b and c show some variation in dose response, but have generally the same shape and orientation as curve d. Note that curve d uses the first *eve* stripe as a landmark, so that curves b and c are expected to be horizontally displaced from d.

Figure 1, Reinitz et al.

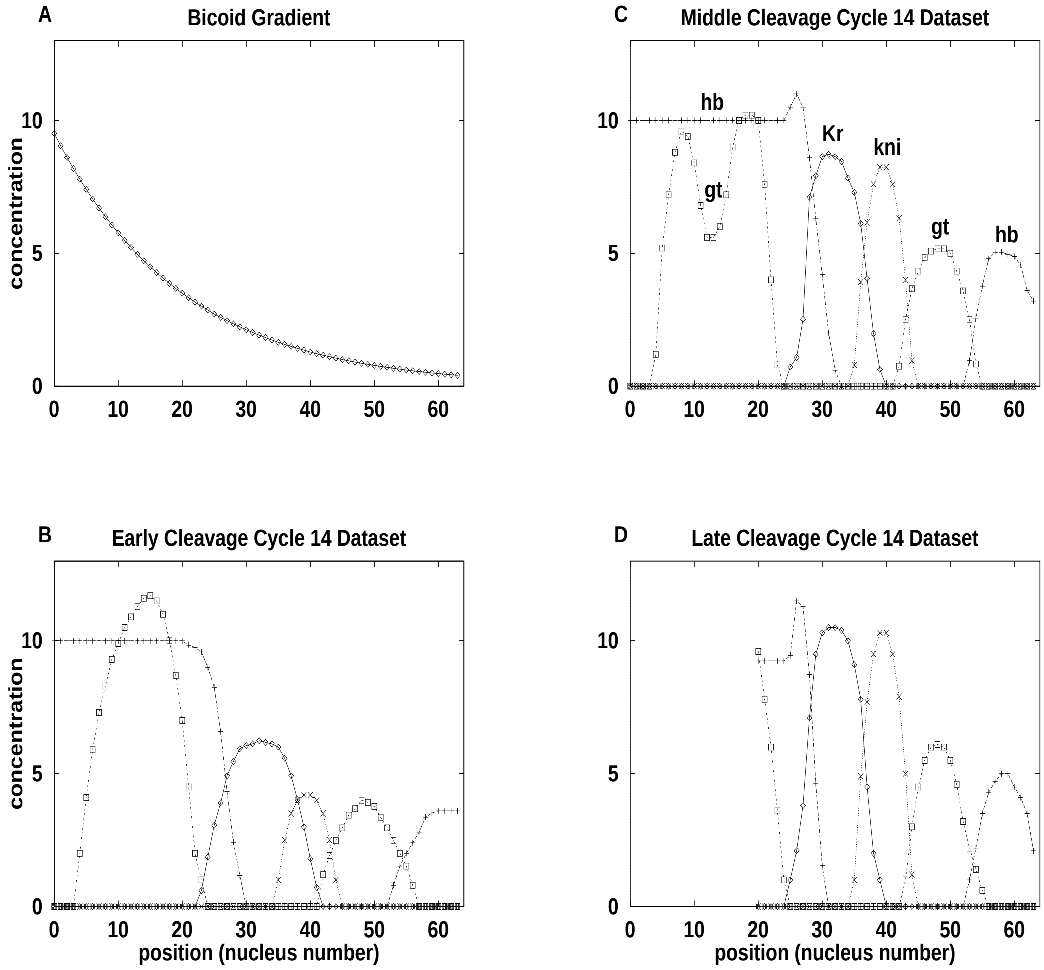

Figure 2, Reinitz et al.

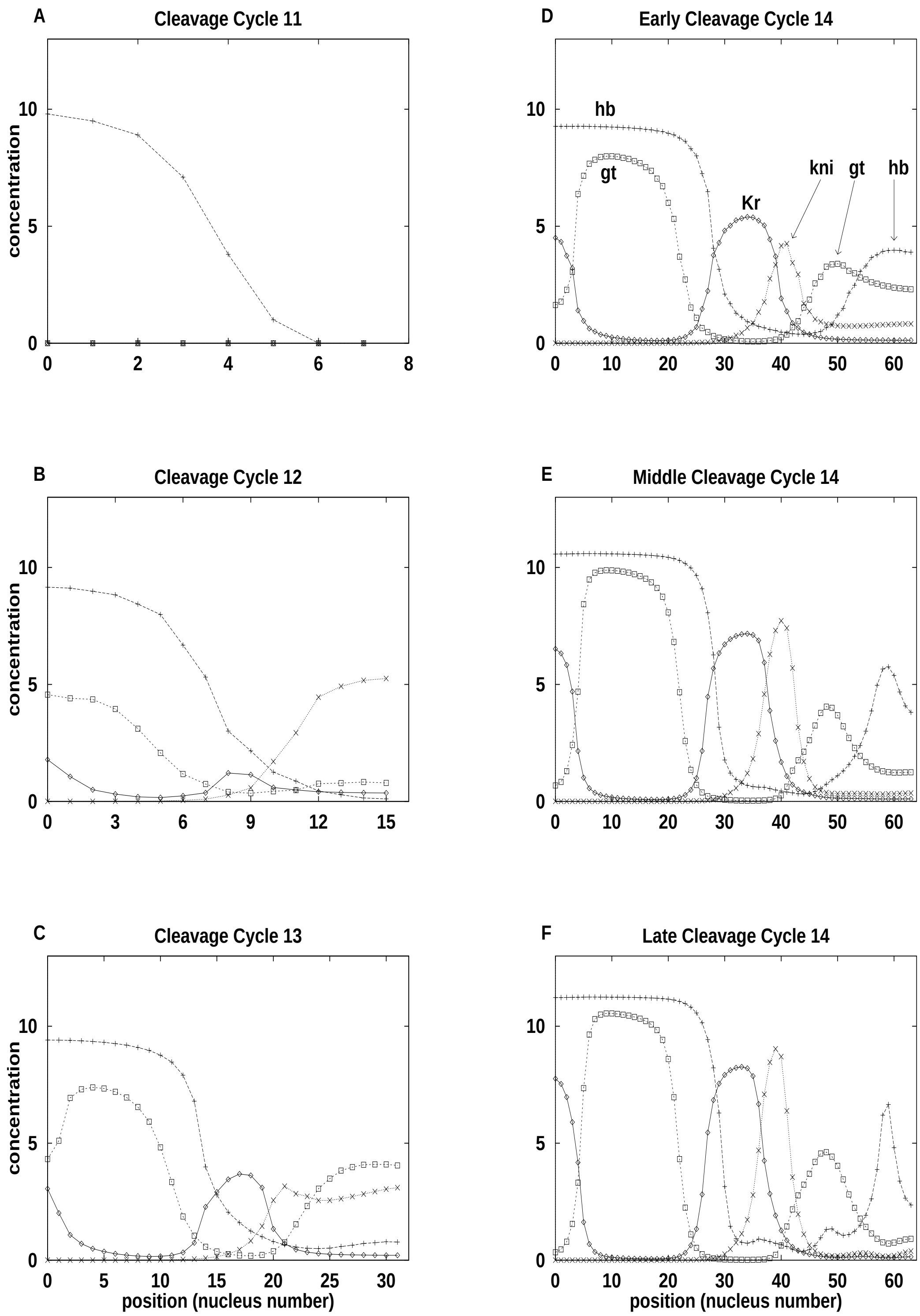

Figure 3, Reinitz et al.

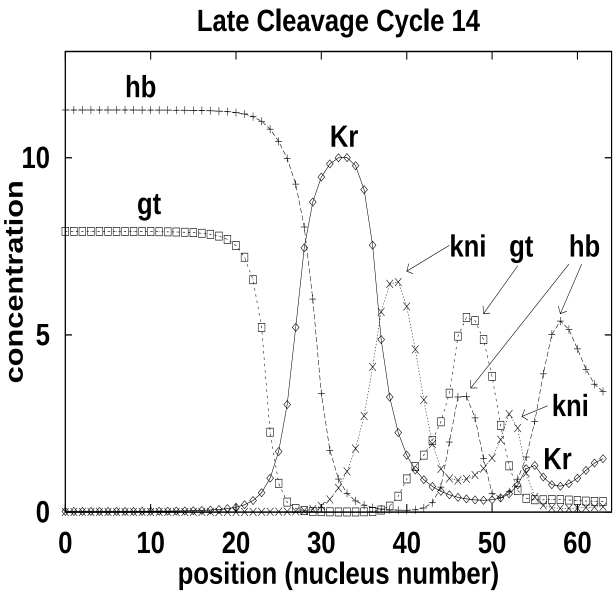

Figure 4, Reinitz et al.