Published in final edited form as:

Sci Signal. ; 4(176): ra38. doi:10.1126/scisignal.2002077.

The cis-Regulatory Logic of Hedgehog Gradient Responses: Key Roles for Gli Binding Affinity, Competition, and Cooperativity

David S. Parker^{1,*†}, Michael A. White^{2,*}, Andrea I. Ramos¹, Barak A. Cohen², and Scott Barolo^{1,‡}

¹Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, MI 48109–2200, USA

²Department of Genetics and Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63108, USA

Abstract

Gradients of diffusible signaling proteins control precise spatial patterns of gene expression in the developing embryo. Here, we use quantitative expression measurements and thermodynamic modeling to uncover the cis-regulatory logic underlying spatially restricted gene expression in a Hedgehog (Hh) gradient in *Drosophila*. When Hh signaling is low, the Hh effector Gli, known as Cubitus interruptus (Ci) in *Drosophila*, acts as a transcriptional repressor; when Hh signaling is high, Gli acts as a transcriptional activator. Counterintuitively and in contrast to previous models of Gli-regulated gene expression, we found that low-affinity binding sites for Ci were required for proper spatial expression of the Hh target gene *decapentaplegic* (*dpp*) in regions of low Hh signal. Three low-affinity Ci sites enabled expression of *dpp* in response to low signal; increasing the affinity of these sites restricted *dpp* expression to regions of maximal signaling. A model

[‡]To whom correspondence should be addressed. sbarolo@umich.edu.

^{*}These authors contributed equally to this work.

[†]Present address: Department of Cell Biology, Duke University Medical Center, Durham, NC 27710, USA.

Author contributions: The experiments were designed by S.B., D.S.P., and A.I.R. and carried out by D.S.P. and A.I.R. B.A.C. and M.A.W. designed and analyzed the thermodynamic models. The paper was written by D.S.P., M.A.W., B.A.C., and S.B.

Competing interests: The authors declare that they have no competing interests.

SUPPLEMENTARY MATERIALS

www.sciencesignaling.org/cgi/content/full/4/176/ra38/DC1

Text 1. Modeling different biophysical mechanisms.

Text 2. Parameter sensitivity analysis of repressor cooperativity and differential affinity models.

Text 3. Testing model fits.

Methods

Fig. S1. Annotated sequences of enhancer constructs.

Fig. S2. Relative in vitro binding affinities of Ci binding sites regulating *ptc* and *dpp*.

Fig. S3. Twelve-species alignments of Ci binding sites.

Fig. S4. Improving the affinity of Ci sites in the *dppD* enhancer causes misinterpretation of Hh gradients and lethal developmental defects.

Fig. S5. Overgrowth and patterning defects in imaginal discs of *dppD-Ci^{Ptc}>dpp* transgenic larvae.

Fig. S6. Best fits of four general classes of biophysical models to high-affinity (3xCi^{Ptc}) and low-affinity (3xCi^{Wt}) data.

Fig. S7. Best-fit equal-affinity models of individual cooperative mechanisms.

Fig. S8. Prediction of single high-affinity site expression by repressor cooperativity and differential affinity models trained on new or old data.

Fig. S9. Repressor cooperativity and differential affinity model sensitivity to parameter changes.

Fig. S10. Repressor and activator occupancy of high- and low-affinity enhancers.

Table S1. Sequences of EMSA oligonucleotides.

Table S2. Parameters included in the model.

Table S3. Best-fit parameter values of four general models.

Table S4. Parameters for repressor cooperativity and differential affinity models.

References

incorporating cooperative repression by Ci correctly predicted the in vivo expression of a reporter gene controlled by a single Ci site. Our work clarifies how transcriptional activators and repressors, competing for common binding sites, can transmit positional information to the genome. It also provides an explanation for the widespread presence of conserved, nonconsensus Gli binding sites in Hh target genes.

INTRODUCTION

Understanding how positional information in morphogen gradients is read requires decoding the interactions between morphogen-responsive transcription factors and the cis-regulatory elements of target genes (1, 2). The cis-regulatory content of the *decapentaplegic* (*dpp*) and *patched* (*ptc*) genes is inconsistent with the currently prevailing model of the Hedgehog (Hh) response in the anterior compartment of the *Drosophila* wing imaginal disc. The expression of these genes has been proposed to depend on an activator threshold mechanism in which transcriptional activators turn on different genes at successively higher amounts of Hh signal (3, 4). The expression of *ptc* is restricted to the anterior/posterior boundary region of highest Hh signal, whereas *dpp* is expressed more broadly within the Hh gradient (Fig. 1A). *ptc* and *dpp* are each regulated by an enhancer containing three binding sites for the Hh-responsive, Gli family transcription factor Cubitus interruptus (Ci) (3, 5, 6). In the presence of Hh, Ci is converted to an activator, but in the absence of Hh, Ci is cleaved to become a transcriptional repressor (7, 8). Thus, the gradient of Hh signal results in opposing gradients of activator and repressor Ci competing for the same genomic binding sites (7–12). In the activator threshold model, the induction of *ptc* and *dpp* is controlled by the abundance of Ci activator in the absence of repressor. *ptc* is induced by high amounts of Ci activator, whereas *dpp* requires less activator for induction (3, 4). As others have argued, this model has implications for the expected affinity of the Ci/Gli binding sites in the enhancers of Hh target genes (13). Genes that respond broadly across the Hh gradient, such as *dpp*, should have high-affinity Ci binding sites, whereas genes activated only by strong Hh signal, such as *ptc*, should have low-affinity sites.

RESULTS

We observed that the enhancers of *dpp* and *ptc* exhibit a regulatory logic opposite that predicted by the activator threshold model. *ptc* is regulated by Ci sites that match the optimal binding sequence (GACCACCCA), whereas *dpp* is regulated by nonconsensus sites of low predicted affinity (Fig. 1, A and B, and fig. S1) (5, 9). We used competitive electrophoretic mobility shift assays (EMSAs) to measure the relative in vitro affinities of Ci sites in the *ptc* and *dpp* enhancers and found that Ci^{ptc} sites in the *ptc* enhancer have considerably higher affinity than Ci^{dpp} sites (Fig. 1B and fig. S2). The predicted superior affinity of Ci^{ptc} sites, relative to Ci^{dpp} sites, is conserved across 12 *Drosophila* species (Fig. 1C and fig. S3). Thus, the regulation of *dpp* and *ptc* in the wing is opposite to that predicted by a simple activator threshold model. *ptc*, which is restricted to the region of highest Hh signal, is regulated by high-affinity sites. In contrast, *dpp*, which responds more broadly in a zone of lower Hh signaling, is regulated by low-affinity sites (Fig. 1A).

To investigate the developmental role of the low-affinity sites in the *dpp* enhancer, we altered all three sites to match the high-affinity Ci binding sequence found in the *ptc* enhancer, a change of only seven nucleotides. We then created transgenic lines containing an extra Flp-inducible copy of *dpp*, driven by the *dpp* disc (*dppD*) enhancer containing either wild-type low-affinity (Ci^{wt}) or altered high-affinity (Ci^{ptc}) sites. An extra copy of *dpp* driven by the low-affinity *dppD*- Ci^{wt} enhancer had no effect on development or survival, whereas the high-affinity Ci^{ptc} enhancer caused lethal developmental defects that

resemble the effects of *dpp* mis-expression in imaginal discs (14), including severe head and limb deformities and pupal lethality resulting from overgrowth fusion, and patterning defects in antenna and leg discs (figs. S4 and S5). These results indicate that the conserved low affinity of the *dppD* enhancer for Ci is functionally relevant.

We developed a quantitative reporter gene assay to further explore the role of low-affinity Ci binding sites in the *dpp* wing disc enhancer. We constructed transgenic fly lines carrying two reporter genes: *dppD*-Ci^{ptc}-RFP, consisting of the high-affinity version of the *dpp* enhancer driving expression of a red fluorescent protein (RFP), and one of several *dpp* enhancers driving green fluorescent protein (GFP) (Fig. 2, A and B). By measuring GFP fluorescence across a transect of the wing pouch and normalizing to peak RFP expression in each disc, we obtained a quantitative readout of both the position and the intensity of GFP reporter activity (Fig. 2B).

Using this assay, we compared the activity of different versions of the *dppD* enhancer driving GFP expression, containing either three low-affinity sites (*dppD*-Ci^{wt}) or three high-affinity sites (*dppD*-Ci^{ptc}) (Fig. 2, A and B). We also measured Ci-independent, “basal” expression from a construct (*dppD*-Ci^{KO}) in which all three Ci sites were mutated to abolish Ci binding (figs. S1 and S2) (9, 15). This basal expression captures the effects of all factors other than Ci on *dpp*, including Engrailed, which directly represses *dpp* near the anterior/posterior boundary (Fig. 1A) (9). This basal construct enabled us to directly measure both activation and repression by Ci, by comparing the activity of the low- or high-affinity enhancers against that of *dppD*-Ci^{KO}. The results show that the response of *dpp* to Hh cannot be explained by an activator threshold model. High-affinity Ci^{ptc} sites caused a posterior shift in stripe position toward the region of strongest Hh signal, whereas low-affinity Ci^{wt} sites produced stronger activation in regions of moderate Hh signal (Fig. 2C). When the basal *dppD*-Ci^{KO}-GFP expression was subtracted from that of *dppD*-Ci^{ptc}-GFP and *dppD*-Ci^{wt}-GFP, we observed that within the zone of moderate Hh signal, low-affinity sites produced activation, whereas high-affinity sites conferred repression (Fig. 2D). This observation shows that Ci_{REP} plays a substantial role in the response to moderate Hh signal, a finding that directly contradicts the assumptions of the activator threshold model (3, 4). Thus, an alternate biophysical model is required to explain the regulatory logic of the *dpp* response to Hh.

To explore biochemical mechanisms that might explain the counterintuitive regulatory logic governing the *dpp* response to Hh, we built a statistical thermodynamic model (16–22) of the *dpp* wing disc enhancer (Fig. 3A). The purpose of the model was to provide a formal framework to explain how particular biophysical interactions might give rise to the *dpp* expression pattern. In the model, gene expression was taken as proportional to the probability of the basal transcription complex (BTC) binding to the target gene. For our purposes, BTC represented a factor whose recruitment is necessary and rate-limiting for transcription and was not meant to stand for any specific molecule or complex. The binding of Ci_{ACT} or Ci_{REP} to the three low- or high-affinity Ci sites in the *dppD* enhancer made the binding of BTC either more or less favorable. The model had free parameters that described the interactions of Ci_{ACT}, Ci_{REP}, and polymerase with DNA and with each other. The model also had two free parameters that determined the shape of the Hh gradient (table S2). We determined the basal transcription rate (namely, the interaction of BTC with the *dpp* promoter in the absence of Ci) by measuring the expression of the *dppD*-Ci^{KO}-GFP enhancer. As described above, the basal expression incorporated the effects of all other regulators and was the baseline against which Ci-mediated activation and repression were defined. This enabled us to focus on the modular effects of Ci on *dpp* while still accounting for the complex effects of other regulators. Because the parameters in this model had clear physical interpretations, and because we explicitly modeled every possible occupancy state

of the *dpp* enhancer, we could use this model to efficiently explore different biophysical mechanisms that might explain the expression pattern of *dpp*.

Different classes of biophysical models are possible within our thermodynamic framework (Fig. 3A and Supplementary Materials, section 1). These include cooperative or anticooperative interactions between repressors, between activators, or between activators and repressors; noncooperative models; models in which activators and repressors exhibit equal affinities for Ci sites; and differential affinity models in which activators and repressors exhibit different affinities for binding sites. We attempted to fit each of these models simultaneously to the low- and high-affinity enhancer–GFP data (fig. S6). Models had between 6 and 11 parameters (tables S2 and S3) and were fit to 514 data points corresponding to expression measurements along the wing disc for both the high- and the low-affinity reporters.

Two biophysical models reproduced the key qualitative feature of the data, which was the zone of moderate Hh signal in which three low-affinity sites cause activation and three high-affinity sites cause repression (Fig. 3, B and C; Supplementary Materials, section 1; and figs. S6 and S7). These models also fit the data well quantitatively (measured by the root mean square of the residuals). The first was a repressor cooperativity model (Fig. 3B and table S4), the key feature of which is that Ci_{REP} , but not Ci_{ACT} , interacts cooperatively with itself. This mechanism is plausible because the cleavage step to produce Ci_{REP} could affect homophilic interactions, either directly or through co-repressors. A noncooperative differential affinity model also fit the data well. In this model, Ci_{ACT} binds better to low-affinity sites than Ci_{REP} , and Ci_{REP} binds better than Ci_{ACT} to high-affinity sites (Fig. 3C and table S4). Both the repressor co-operativity and differential affinity models successfully predicted the posterior shift of the anterior boundary of activation from the *dppD*- Ci^{ptc} enhancer (Fig. 3, B and C; activation boundary indicated by circles). Thus, both models recapitulated the most distinct feature of the enhancer–GFP data, which is that under moderate Hh signal, low-affinity sites produce activation, whereas high-affinity sites produce repression. In other respects, the differential affinity model performed less well, particularly in the anterior region where *dppD*- Ci^{wt} -GFP exhibits repression (Fig. 3C). However, because both models captured the most important qualitative feature and fit the data well quantitatively, we considered both models strong candidate biophysical scenarios that could explain the *dpp* response to Hh.

We sought to distinguish between the repressor cooperativity and differential affinity models experimentally. We found that the best-fit differential affinity and repressor cooperativity models made different predictions about the behavior of a *dpp* enhancer with only a single high-affinity Ci binding site. Because cooperativity between repressors should not occur at a lone Ci site, the repressor cooperativity model predicted that a single high-affinity site will produce activation at moderate Hh signal, and three high-affinity sites will cause repression. Indeed, the repressor cooperativity model predicted that a single, high-affinity site will produce a broader region of activation than the wild-type enhancer (Fig. 4A). By contrast, the differential affinity model predicted that the zones of activation and repression by a single high-affinity site construct will exactly coincide with the regions of activation and repression produced by three high-affinity sites (Fig. 4A). This is because zones of repression and activation depend on the ratio of bound Ci_{ACT} to bound Ci_{REP} ; because the differential affinity model is noncooperative, the ratio of bound Ci_{ACT} to bound Ci_{REP} is not affected by the number of binding sites.

To discriminate between the repressor cooperativity and differential affinity models in vivo, we constructed a transgenic fly line carrying a GFP reporter with a single high-affinity Ci site. A basal amount of gene expression across the wing disc was created by adding three

binding sites for the Hh-independent transcriptional activator Grainyhead (23), next to a single Ci^{ptc} site (fig. S1, C and D). The observed expression pattern supports the repressor cooperativity model over the differential affinity model (Fig. 4A), despite the fact that our original models were not adjusted to account for the effect of the Grainyhead sites. Refitting the models to the new data produces similar qualitative results (fig. S8). Consistent with the predictions of the repressor cooperativity model, a single high-affinity site produced a broad but attenuated stripe with an anterior boundary that extends past the boundary of the wild-type stripe. Thus, in cells receiving low-to-moderate Hh signaling, three high-affinity sites confer repression, whereas a single high-affinity site confers activation. This gradient response profile rules out the differential affinity model and could not have been predicted from the activator threshold model. The repressor cooperativity model also successfully predicted the relative affinities of the high- and low-affinity sites. The average of the ratios of the measured relative affinities of the three high-affinity and the three low-affinity sites was 40-fold (Fig. 1B). The ratio of the high- and low-affinity Ci -DNA association constants in the repressor cooperativity model, a ratio that was free to vary during the fitting procedure, was 50-fold, close to the measured ratio (table S4). Thus, the repressor cooperativity model accurately predicts the transcriptional response to Hh under various conditions and also predicts the relative affinities of the low- and high-affinity sites for Ci . The repressor cooperativity model therefore best explains the positioning of *dpp* and *ptc* in the Hh gradient.

DISCUSSION

We have shown that spatial information in the wing disc Hh gradient is interpreted by a cis-regulatory logic that relies on activator-repressor competition, which is modulated by binding site affinity and cooperative repression. In previous studies of Hh target genes, the role of the affinity of the Gli or Ci binding site has been neglected or has been assumed to play a role opposite to what our data show (3–5, 8, 13). Moreover, the currently accepted activator threshold model of the transcriptional response to Hh assumes that the role of Gli and Ci repressors is limited to regions of little or no Hh signal (Fig. 4B) (3, 4, 13). No previously described model of Hh response, including the activator threshold model, can account for the observations we describe here. Our data show that substantial repression can occur even at moderate Hh signal (Fig. 2D) and suggest that the transcriptional response in much of the Hh gradient depends on the outcome of a competition between Ci_{ACT} and Ci_{REP} for enhancer binding (Fig. 4B). This new model of the cis-regulatory logic underlying Hh response integrates the effects of both Ci_{ACT} and Ci_{REP} along the entire Hh gradient and explains the importance of low-affinity Gli binding sites in the positioning of gene expression.

Our results suggest that the low affinity of the *dpp* enhancer for Ci can be explained by the need to mitigate the effects of cooperative repression in a region of the gradient where substantial amounts of Ci_{REP} are present, while still allowing activation by Ci_{ACT} . Within the context of our model, repressor cooperativity is defined as any interaction that makes the binding of additional Ci_{REP} more favorable when one Ci_{REP} is already bound. Cooperativity could arise from direct interactions between Ci_{REP} or from interactions of Ci_{REP} with other transcription factors, cofactors, or histones. In principle, Ci_{REP} cooperativity at the enhancer could be attenuated by various cis-regulatory strategies besides lowering binding affinity, such as reducing the number of Ci sites (Fig. 4A) or increasing their spacing. However, these alternative strategies may not be equally able to maintain activation by Ci_{ACT} in regions of low Hh signaling.

The posterior-to-anterior gradient of Hh in the wing disc establishes opposing gradients of Ci_{ACT} and Ci_{REP} . High amounts of Ci_{ACT} are present at the anterior/posterior boundary,

whereas more anterior regions feature high amounts of Ci_{REP}. Within the intermediate zone of the gradient, mixed amounts of Ci_{ACT} and Ci_{REP} compete for enhancer binding, with activation and repression determined by the ratio of bound Ci_{ACT} to bound Ci_{REP} (Fig. 4B). In the repressor cooperativity model, repressors outcompete activators for binding at high-affinity enhancers, but not at low-affinity enhancers, in this region of the gradient. Several morphogen signaling pathways have the potential to produce reciprocal gradients of repressors and activators competing for common binding sites (1, 2). We have presented the first detailed mechanistic model that explains how reciprocal gradients of Gli activators and repressors are transcriptionally interpreted. A similar regulatory logic may inform responses to other morphogens that control transcriptional switches, particularly those whose target genes are regulated by low-affinity sites (9–12, 24–26).

Our model provides an explanation for the widespread presence of evolutionarily conserved, nonconsensus Gli or Ci binding sites in the enhancers of Hh target genes. With the exception of *ptc*, all known direct targets of Hh in *Drosophila* are regulated by nonconsensus Ci sites with predicted low affinity (9, 14, 25, 26). Our results indicate that low-affinity sites are necessary to position a stripe of expression in the middle of a Hh gradient, because high-affinity sites induce repression outside the zone of strongest signaling. Because mammalian Gli target genes are also regulated by nonconsensus sites (10, 11), our conclusions may also apply to vertebrate Hh targets: Such targets may acquire weak-affinity Gli sites to minimize cooperative repression in a Gli cross-gradient.

In a few documented cases, low-affinity transcription factor binding sites have important spatiotemporal patterning functions. Two well-studied examples are the response to the morphogen Dorsal in *Drosophila* and the temporal control of developmental gene expression (27–29). In these cases, low-affinity sites set a high threshold for activator concentration, restricting activation to cells or times in which activator concentration is maximal. In the case described here, an opposite cis-regulatory logic applies: Low-affinity sites are specifically required for activation in cells receiving lower amounts of signal. This is a consequence of the fact that in this region of the gradient, activators and repressors compete for the same genomic binding sites.

Our results suggest that most current biochemical and computational approaches to identifying Hh target genes, which typically focus on the highest-affinity Ci or Gli sites (10–12), may overlook a large proportion of important Hh target genes. More generally, transcriptional cooperativity may play an important cis-regulatory role in enhancers with conserved low-affinity binding sites (27, 29, 30).

MATERIALS AND METHODS

EMSA competition assays

Ci protein was made in vitro with the TNT SP6 High-Yield Wheat Germ Protein Expression System (Promega). A polymerase chain reaction (PCR)-amplified segment of Ci complementary DNA (cDNA), encoding a region of the protein including the DNA binding domain (residues 441 to 674), was used as a template. Primer sequences were as follows (underlined bases denote portions of primers corresponding to Ci cDNA sequence): 5'-AGATCAGATTTATTTAGGTGACACTATAGAACA-GACCACCATGATCAAGGATGAACCCGGAGATTTCA-3' and 5'-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATCACTAAGTGATGGTTTGATGAGCTTATATC-3'. Oligonucleotides corresponding to Ci site 1 of the *patched* enhancer (*ptc1*; see table S1 for oligonucleotide sequences) were end-labeled with [γ -³²P]ATP (adenosine triphosphate) using T4 polynucleotide kinase, annealed, and used as the hot probe. Radiolabeled probe (1.6 nmol) and various amounts of unlabeled competitor

oligonucleotides (where appropriate; see below) were mixed and incubated on ice for 15 min with Ci protein under the following conditions: 10 mM tris-HCl (pH 7.5), 50 mM NaCl, 10 mM dithiothreitol, 1 mM EDTA, salmon sperm DNA (27.5 µg/ml), and poly(dI:dC) (polydeoxyinosine-polydeoxycytosine; 100 µg/ml). Reactions were analyzed by electrophoresis on a 4% nondenaturing polyacrylamide gel in 0.5× TBE (tris-borate EDTA) buffer. After drying, gels were exposed to a Storage Phosphor screen (GE) overnight and the data were collected with a Typhoon Variable Mode Imager (GE).

Bands were quantified with ImageJ software (W. S. Rasband, U.S. National Institutes of Health, Bethesda, MD; <http://rsb.info.nih.gov/ij>). To determine relative binding affinities, we first measured the amount of bound radiolabeled ptc1 probe (as a percentage of free probe) when no competitor was present. We then measured the decrease in bound radiolabeled ptc1 upon addition of increasing amounts of nonradiolabeled competitor oligonucleotides (fig. S3). From these data, we calculated the fold excess of nonradiolabeled oligonucleotide required to compete away 30% of bound ptc1. Affinities in Fig. 1C are relative to that of ptc1.

DNA cloning and mutagenesis

Enhancer constructs were TOPO-cloned into the pENTR/DTOPO plasmid (Invitrogen) and then cloned into the P-element transformation vectors pH-Stinger (31) (for full-length enhancer constructs) or pH-Red Stinger (32) (for RFP constructs). Mutated enhancers and *dpp* expression constructs were created with a combination of overlap extension PCR and assembly PCR (33). Detailed protocols are available at <http://sitemaker.umich.edu/barolo/protocols>. Annotated enhancer sequences are given in fig. S1.

Drosophila transgenic strains

P-element transformation by embryo injection was performed essentially as previously described (34). An updated protocol is available at <http://sitemaker.umich.edu/barolo/protocols>. *w¹¹¹⁸* flies were used for transgenesis. To eliminate the effects of genomic insertion site on gene expression, we examined several independent transgenic lines bearing each construct (fig. S9); representative results are shown. For experiments in Fig. 4, a phiC31 landing site was used to eliminate position effects. P[*HS-Flp*] was a gift of K. Cadigan.

Whole-mount staining and microscopy

For collecting fluorescence data, imaginal discs from wandering third instar larvae were fixed in 4% paraformaldehyde in phosphate-buffered saline. Discs were mounted in ProLong Gold with 4',6-diamidino-2-phenylindole (DAPI) (Invitrogen). Fluorescence images were captured on an Olympus FluoView 500 Laser Scanning Confocal Microscope mounted on an Olympus IX-71 inverted microscope. Discs were imaged under identical confocal microscopy conditions and settings.

Binding site consensus scores

Matrix similarity scores were calculated as described (35) with in vitro Ci binding data generated by Hallikas and colleagues (12).

Quantitation of transgenic reporter expression data

In all fluorescence quantitation experiments, each sample was processed identically, and identical confocal settings were used for all experiments from which results were directly compared. Fluorescence data were quantitated with ImageJ software. For each construct,

data were taken from a transgenic line with the median total GFP expression; other transgenic lines gave comparable curves (fig. S9). For experiments in Fig. 4, a phiC31 landing site was used to eliminate position effects. For wing imaginal disc experiments in Figs. 2 and 4, data were collected from a rectangle centered on the dorsal half of the wing pouch (black rectangle in Fig. 2B). Anterior/posterior position was calculated as a percentage of disc width. For each wing disc, GFP intensity was normalized to peak *dppD*-Ci^{ptc}-RFP intensity. The position of peak *dppD*-Ci^{ptc}-RFP expression was used to align GFP data from different discs.

Thermodynamic model

We implemented an occupancy model of the *dppD* enhancer based on the framework developed by Shea and Ackers (16) and Buchler *et al.* (17) as described previously (19). We modeled binding of activator (Ci_{ACT}) and repressor (Ci_{REP}) to three identical low- or high-affinity Ci binding sites, and the binding of the BTC to a single BTC “site.” In this model, BTC could be any factor or complex whose recruitment is necessary and rate-limiting for target gene transcription and that interacts with Ci proteins bound to any Ci site on the enhancer. This factor could be polymerase, a critical coactivator, or a chromatin remodeling activity, for example. Gene expression was taken as proportional to the probability of BTC binding to the *dpp* promoter. The probability of BTC binding was determined by BTC interaction with the promoter and by interactions of Ci_{ACT} and Ci_{REP} with DNA, with each other, and with BTC. To calculate the probability of BTC binding, we first calculated a statistical weight W_k for each possible occupancy state k of the *dpp* enhancer as follows:

$$W_k = \exp \left(- \sum_i q_i - \sum_j \omega_{ij} \right) \quad (1)$$

where the summation was over the protein-DNA interaction Gibbs free energies q for each occupied Ci or BTC binding site i , and each protein-protein interaction Gibbs free energy ω (between neighboring occupied sites i and j , or between Ci bound at site i and BTC bound to its respective site; see Fig. 3A). Next, we summed the statistical weights of the BTC-bound occupancy states b to obtain the relative probability of BTC binding. We normalized this relative probability by dividing by the sum of the weights of all enhancer states k to obtain the absolute probability of BTC binding, which was proportional to gene expression:

$$\text{Expression} = \alpha \left(\frac{\sum W_b}{\sum W_k} \right) \quad (2)$$

For this study, we let the proportionality constant α equal 1, because the transgenic reporter gene measurements as performed were normalized to fall between 0 and 1.

Ci-DNA interaction energies q were calculated from Ci concentrations and Ci-DNA association constants K as follows:

$$q = - \ln K - \ln [\text{Ci}] \quad (3)$$

Basal, Ci-independent transcription was determined by the BTC-DNA interaction energy (q_{BTC}), which was calculated for each position in the gradient directly from the *dppD*-Ci^{KO} transgenic reporter data as follows:

$$q_{BTC} = \ln(1 - \text{Expression}_{Ci^{KO}}) - \ln(\text{Expression}_{Ci^{KO}}) \quad (4)$$

Equation 4 was a rearrangement of the thermodynamic model applied only to BTC binding:

$$\text{Expression}_{Ci^{KO}} = \frac{e^{-q_{BTC}}}{1 + e^{-q_{BTC}}} \quad (5)$$

Basal transcription thus varied across the gradient and accounted for the effects of all factors (other than Ci) that regulate *dpp* in the wing disc.

We did not model Hh signaling. The Hh gradient was implemented as a gradient of Ci_{ACT} , which reached maximum concentration at the anterior/posterior boundary and decreased exponentially in the anterior direction:

$$Ci_{ACT} = h \times \exp\left(-\frac{x}{\sqrt{D}}\right) \quad (6)$$

where x is the distance from the anterior/posterior boundary, D determines the steepness of the gradient, and h scales the Ci_{ACT} concentration values.

Total Ci (Ci_{TOT}) was the total concentration of Ci in any form at each position in the gradient. For simplicity, Ci_{TOT} was kept constant across the gradient, although this assumption was not crucial for the model. The model still succeeded if Ci_{TOT} decreased with distance from the anterior/posterior boundary; as long as Ci_{TOT} decreased slowly with respect to the gradient of Ci_{ACT} , our model results still hold. At the anterior/posterior boundary, x_{min} , all available Ci was completely in the activator form, and, therefore, the Ci_{TOT} was calculated from Eq. 6 at x_{min} :

$$Ci_{TOT} = h \times \exp\left(-\frac{x_{min}}{\sqrt{D}}\right) \quad (7)$$

Any Ci not in activator form was in repressor form:

$$Ci_{REP} = Ci_{TOT} - Ci_{ACT} \quad (8)$$

Thus, in our model, repressors could be present in areas of intermediate Hh signal, in a manner that depends on gradient shape determined by parameter D .

In equal-affinity models, Ci_{ACT} and Ci_{REP} shared the same association constants for each class of binding site, K_{wt} (low affinity) and K_{ptc} (high affinity). In differential affinity models, there were association constants for each class of site and each form of Ci, $K_{wt-activator}$, $K_{wt-repressor}$, $K_{ptc-activator}$, and $K_{ptc-repressor}$.

In summary, at each position in the gradient, there were specific values of Ci_{ACT} , Ci_{REP} , and q_{BTC} , which were determined by the gradient exponential function and by the *dppD-Ci^{KO}* data. The model took as input position in the gradient and calculated the probability of BTC binding at that position. The model was applied at each of 257 positions in the gradient for both high- and low-affinity enhancers to generate the predicted expression pattern of *dpp* in the wing disc.

We fit between 6 and 11 parameters to 514 transgenic reporter gene measurements from a single experiment (257 measurements each of the *dppD*-Ci^{wt} and *dppD*-Ci^{ptc} enhancers). The model parameters are shown in table S2.

Fits were performed with the trust-region-reflect algorithm of the *lsqcurvefit* function in MATLAB version 2010b (The Mathworks). To set the appropriate scale for the model parameters, the high-affinity association constant (for equal-affinity models) or the high-affinity, Ci_{ACT} association constant (for the differential affinity model) was held fixed at an arbitrarily chosen but biologically reasonable value of 5×10^8 , whereas all other parameters were fit. Our initial parameter guesses were chosen to take biologically reasonable values, but randomly varied to cover a broad range of initial values (for example, association constants could range over six orders of magnitude, *h* could range from nanomolar to millimolar concentration, and so on). For each model, we performed at least 200 different fits and found that the best ~50 fits (determined by the root mean square of the residuals) always converged on similar parameter values, indicating that the fitting algorithm reproducibly converged on a single least-squares minimum within the allowable range of parameter values.

Different biophysical mechanisms were modeled by allowing different Ci-Ci interaction parameters (ω) to be either negative (cooperative), positive (anticooperative), or zero (noncooperative). For example, in the repressor cooperativity model, $\omega(\text{Ci}_{\text{REP}}\text{-Ci}_{\text{REP}})$ was fit to a negative value, whereas $\omega(\text{Ci}_{\text{ACT}}\text{-Ci}_{\text{ACT}})$ and $\omega(\text{Ci}_{\text{ACT}}\text{-Ci}_{\text{REP}})$ were fixed at zero. MATLAB code for the thermodynamic model is available upon request by contacting cohen@genetics.wustl.edu.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank R. Mitra, M. Brent, D. Wellik, T. Glaser, D. Gumucio, and members of our labs for helpful discussions and comments on the manuscript.

Funding: This work was supported by grants from the NIH to B.A.C. and S.B., including GM078222 and GM076509 and American Recovery and Reinvestment Act supplements GM07822203S1 and GM07650903S1. A.I.R. was supported by the University of Michigan Cellular and Molecular Biology Training Program.

REFERENCES AND NOTES

1. Barolo S, Posakony JW. Three habits of highly effective signaling pathways: Principles of transcriptional control by developmental cell signaling. *Genes Dev.* 2002; 16:1167–1181. [PubMed: 12023297]
2. Ashe HL, Briscoe J. The interpretation of morphogen gradients. *Development.* 2006; 133:385–394. [PubMed: 16410409]
3. Wang QT, Holmgren RA. The subcellular localization and activity of *Drosophila* Cubitus interruptus are regulated at multiple levels. *Development.* 1999; 126:5097–5106. [PubMed: 10529426]
4. Lum L, Beachy PA. The Hedgehog response network: Sensors, switches, and routers. *Science.* 2004; 304:1755–1759. [PubMed: 15205520]
5. Held, LI. Imaginal Discs: The Genetic and Cellular Logic of Pattern Formation. Cambridge University Press; Cambridge, MA: 2002.
6. Alexandre C, Jacinto A, Ingham PW. Transcriptional activation of hedgehog target genes in *Drosophila* is mediated directly by the cubitus interruptus protein, a member of the GLI family of zinc finger DNA-binding proteins. *Genes Dev.* 1996; 10:2003–2013. [PubMed: 8769644]

7. Aza-Blanc P, Ramírez-Weber FA, Laget MP, Schwartz C, Kornberg TB. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. *Cell*. 1997; 89:1043–1053. [PubMed: 9215627]
8. Méthot N, Basler K. Hedgehog controls limb development by regulating the activities of distinct transcriptional activator and repressor forms of Cubitus interruptus. *Cell*. 1999; 96:819–831. [PubMed: 10102270]
9. Müller B, Basler K. The repressor and activator forms of Cubitus interruptus control Hedgehog target genes through common generic Gli-binding sites. *Development*. 2000; 127:2999–3007. [PubMed: 10862738]
10. Vokes SA, Ji H, McCuine S, Tenzen T, Giles S, Zhong S, Longabaugh WJR, Davidson EH, Wong WH, McMahon AP. Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. *Development*. 2007; 134:1977–1989. [PubMed: 17442700]
11. Vokes SA, Ji H, Wong WH, McMahon AP. A genome-scale analysis of the *cis*-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes Dev*. 2008; 22:2651–2663. [PubMed: 18832070]
12. Hallikas O, Palin K, Sinjushina N, Rautiainen R, Partanen J, Ukkonen E, Taipale J. Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell*. 2006; 124:47–59. [PubMed: 16413481]
13. Dessaud E, McMahon AP, Briscoe J. Pattern formation in the vertebrate neural tube: A sonic hedgehog morphogen-regulated transcriptional network. *Development*. 2008; 135:2489–2503. [PubMed: 18621990]
14. Morimura S, Maves L, Chen Y, Hoffmann FM. *decapentaplegic* overexpression affects *Drosophila* wing and leg imaginal disc development and *wingless* expression. *Dev Biol*. 1996; 177:136–151. [PubMed: 8660883]
15. Von Ohlen T, Hooper JE. Hedgehog signaling regulates transcription through Gli/Ci binding sites in the *wingless* enhancer. *Mech Dev*. 1997; 68:149–156. [PubMed: 9431812]
16. Shea MA, Ackers GK. The O_R control system of bacteriophage lambda. A physical-chemical model for gene regulation. *J Mol Biol*. 1985; 181:211–230. [PubMed: 3157005]
17. Buchler NE, Gerland U, Hwa T. On schemes of combinatorial transcription logic. *Proc Natl Acad Sci USA*. 2003; 100:5136–5141. [PubMed: 12702751]
18. Segal E, Raveh-Sadka T, Schroeder M, Unnerstall U, Gaul U. Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature*. 2008; 451:535–540. [PubMed: 18172436]
19. Gertz J, Siggia ED, Cohen BA. Analysis of combinatorial *cis*-regulation in synthetic and genomic promoters. *Nature*. 2009; 457:215–218. [PubMed: 19029883]
20. He X, Samee MA, Blatti C, Sinha S. Thermodynamics-based models of transcriptional regulation by enhancers: The roles of synergistic activation, cooperative binding and short-range repression. *PLoS Comput Biol*. 2010; 6:e1000935. [PubMed: 20862354]
21. Janssens H, Hou S, Jaeger J, Kim AR, Myasnikova E, Sharp D, Reinitz J. Quantitative and predictive model of transcriptional control of the *Drosophila melanogaster even-skipped* gene. *Nat Genet*. 2006; 38:1159–1165. [PubMed: 16980977]
22. Fakhouri WD, Ay A, Sayal R, Dresch J, Dayringer E, Arnosti DN. Deciphering a transcriptional regulatory code: Modeling short-range repression in the *Drosophila* embryo. *Mol Syst Biol*. 2010; 6:341. [PubMed: 20087339]
23. Furriols M, Bray S. A model Notch response element detects Suppressor of Hairless-dependent molecular switch. *Curr Biol*. 2001; 11:60–64. [PubMed: 11166182]
24. Winklmayr M, Schmid C, Laner-Plamberger S, Kaser A, Aberger F, Eichberger T, Frischauf AM. Non-consensus GLI binding sites in Hedgehog target gene regulation. *BMC Mol Biol*. 2010; 11:2. [PubMed: 20070907]
25. Hersh BM, Carroll SB. Direct regulation of *knot* gene expression by Ultrabithorax and the evolution of *cis*-regulatory elements in *Drosophila*. *Development*. 2005; 132:1567–1577. [PubMed: 15753212]

26. Kwon C, Hays R, Fetting J, Orenic TV. Opposing inputs by Hedgehog and Brinker define a stripe of *hairy* expression in the *Drosophila* leg imaginal disc. *Development*. 2004; 131:2681–2692. [PubMed: 15128656]
27. Jiang J, Levine M. Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell*. 1993; 72:741–752. [PubMed: 8453668]
28. Gaudet J, Mango SE. Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4. *Science*. 2002; 295:821–825. [PubMed: 11823633]
29. Rowan S, Siggers T, Lachke SA, Yue Y, Bulyk ML, Maas RL. Precise temporal control of the eye regulatory gene *Pax6* via enhancer-binding site affinity. *Genes Dev*. 2010; 24:980–985. [PubMed: 20413611]
30. Nguyen V, Chokas AL, Stecca B, Altaba ARi. Cooperative requirement of the Gli proteins in neurogenesis. *Development*. 2005; 132:3267–3279. [PubMed: 15983404]
31. Barolo S, Carver LA, Posakony JW. GFP and β -galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. *BioTechniques*. 2000; 29:726, 728, 730, 732. [PubMed: 11056799]
32. Barolo S, Castro B, Posakony JW. New *Drosophila* transgenic reporters: Insulated P-element vectors expressing fast-maturing RFP. *BioTechniques*. 2004; 36:436–440. [PubMed: 15038159]
33. Johnson LA, Zhao Y, Golden K, Barolo S. Reverse-engineering a transcriptional enhancer: A case study in *Drosophila*. *Tissue Eng Part A*. 2008; 14:1549–1559. [PubMed: 18687053]
34. Rubin GM, Spradling AC. Genetic transformation of *Drosophila* with transposable element vectors. *Science*. 1982; 218:348–353. [PubMed: 6289436]
35. Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: New fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*. 1995; 23:4878–4884. [PubMed: 8532532]

Fig. 1.

The Hh target genes *ptc* and *dpp* differ in sequence and affinity of Ci/Gli binding sites in their enhancers. **(A)** Diagram of the third-instar wing disc, showing individual gene and protein expression patterns across a transect of the disc. The dashed line indicates the anterior/posterior boundary. Engrailed (En) directly represses *dpp* in cells closest to the boundary, as shown by the white bar. **(B)** Sequence and in vitro relative affinity of the three Ci binding sites within enhancers of *ptc* (red) and *dpp* (blue). Matches to the optimal consensus sequence are shown as dashes. **(C)** Matrix similarity scores (predicted affinity) of Ci sites in *ptc* (red) and *dpp* (blue), across 12 *Drosophila* species genomes. *Drosophila melanogaster* scores are shown as black dots.

Fig. 2.

Optimizing the affinity of Ci sites in the *dpp* enhancer restricts expression to the most strongly Hh-responding cells. **(A)** Diagrams of versions of the *dpp* disc enhancer (*dppD*) with low-affinity Ci binding sites [blue, wild type (WT)], high-affinity Ci sites (red, Ci^{ptc}), or mutated sites that do not bind Ci (black, Ci^{KO}). **(B)** Expression of altered *dpp* enhancers across the Hh gradient of the wing imaginal disc. Confocal images of segments of wing discs from double-transgenic reporter larvae are shown. Green, *dppD*-GFP reporters; magenta, *dppD*-Ci^{ptc}-RFP used as an internal control for fluorescence and stripe position. **(C)** Normalized GFP fluorescence data collected from wing discs. Arrows show medians of fluorescence along the anterior-posterior axis for each construct. Yellow circles indicate the positions at which Ci switches from net activation to net repression. **(D)** Net effect of wild-type or high-affinity Ci sites on *dppD* expression (normalized expression minus normalized *dppD*-Ci^{KO} expression). Increased Ci affinity provides stronger activation in region 1, as well as stronger repression in region 3, but switches from activation to repression in region 2. Yellow circles indicate the positions at which Ci switches from net activation to net repression. A, anterior; P, posterior.

Fig. 3.

Thermodynamic model of the *dpp* response to Hh. **(A)** Schematic of the protein-protein and protein-DNA interactions captured by the model. Each arrow represents a free energy (ΔG) that determines the probability of specific interactions occurring on the *dpp* enhancer. **(B)** The repressor cooperativity model accurately reproduces the low-affinity (Ci^{dpp}) and high-affinity (Ci^{ptc}) enhancer-GFP data. Modeled relative expression (enhancer expression minus basal expression, y axis) accurately captures the position of the measured switch from activation to repression (circles) along the Hh gradient (x axis). RMS = root mean square of residuals; vertical line indicates A/P boundary. **(C)** A noncooperative differential affinity model, in which Ci_{ACT} exhibits higher affinity for Ci^{dpp} sites than Ci_{REP}, and Ci_{REP} exhibits higher affinity for Ci^{ptc} sites than Ci_{ACT}, also accurately models the switch from activation to repression.

Fig. 4.

The repressor cooperativity model successfully predicts expression from a single high-affinity Ci site. **(A)** Relative measured GFP expression from a $1xCi^{ptc}$ enhancer compared with $1xCi^{ptc}$ predictions of the repressor cooperativity (RC) and differential affinity (DA) models. $3xCi^{dpp}$ and $3xCi^{ptc}$ are the original data sets on which the models were trained. The vertical line indicates the A/P boundary. **(B)** A model incorporating cooperative repression and repressor-activator competition explains why intermediate amounts of Hh produce activation from low-affinity sites and repression from high-affinity sites. The Hh gradient produces opposing gradients of Ci^{ACT} and Ci^{REP} competing for common binding sites. At intermediate amounts of Ci^{ACT} and Ci^{REP} , the high-affinity enhancer, which has higher overall occupancy than the low-affinity enhancer, favors cooperative repression by Ci^{REP} . In each example, only one of many possible binding states is shown.