Function of synthetic Hox transcription factors *in vivo* and the quantitative study of their molecular interactions with nuclear DNA at the single-molecule level in live cells

Inauguraldissertation

Zur Erlangung der Würde eines Doktors der Philosophie vorgelegt der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel von

Dimitrios Papadopoulos

aus Marousi, Griechenland

Basel, 2010

creative commons

Attribution-Noncommercial-No Derivative Works 2.5 Switzerland

---

You are free:

to Share — to copy, distribute and transmit the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- Nothing in this license impairs or restricts the author's moral rights.

Your fair dealing and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license) available in German:
http://creativecommons.org/licenses/by-nc-nd/2.5/ch/legalcode.de

Disclaimer:
The Commons Deed is not a license. It is simply a handy reference for understanding the Legal Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license. Creative Commons is not a law firm and does not provide legal services. Distributing of, displaying of, or linking to this Commons Deed does not create an attorney-client relationship.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/deed.en Datum: 3.4.2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag der Herren

Prof. Dr. Walter J. Gehring

Referent

Prof. Dr. Rudolf Rigler

Koreferent

Basel, den 25.05.2010

Prof. Dr. Eberhard Parlow

Dekan

der Philosophisch-Naturwissenschaftlichen Fakultät

Universität Basel

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to express my gratitude to Prof. W. J. Gehring for our whole scientific interaction, his support, his overwhelming enthusiasm, his trust and great expectations of me, elements that filled me with respect and admiration throughout the years of my thesis. He has been a true mentor to me, having taught me not only genetics, but also how to understand and respect science, and, above all, enjoy its beauty. His guidance has always been on my side in the most peculiar way. By setting the problems and allowing me to “incubate” them in my mind, fall and rise, he has stimulated and enhanced over the years my scientific thinking. He has also gratefully provided his advice and helped me address the questions and find the answers in the most astute manner. Above all, however, I have admired him as a person for being enthusiastic as a teenager but wise as a grandfather, loving as a friend but fair as a judge, cautious as a scientist but protective as a tutor. All these qualities have made me easily jump out of bed in the morning, skip lunch at noon and drive back home singing in the night, but at the same time have filled me with pride and the eagerness to live up to all expectations. I realize now that these merits are not given to all people by default and therefore I will endeavor to maintain the virtues that I have acquired and to transmit them to whoever I can in the future.

I would also like to thank Prof. R. Rigler for our excellent collaboration, his support and enthusiasm during the realization of such a challenging project. He has granted to me equally as much zeal and has believed in me whole-heartedly, kindly providing his deep experience and offering his help at all possible times. He has also been extremely patient and encouraging through the difficult times of the project and stood always on my side, suggesting, explaining, and constantly providing ideas and solutions to all sorts of upcoming problems. His attitude has resulted in all us being ultimately productive and I am really looking forward to all that is to follow next.

No less gratitude for their outstanding contribution and their friendship is attributed to Prof. D. Reséndez-Pérez and her student D. L. Cárdenas-Chávez. They have actively participated in our “*Antennapedia*” project and have facilitated greatly the production of a large amount of results in a relatively short period of time. Without their contribution this project would have been impossible to complete as effectively. Our collaboration and positive interactions have been extraordinary.

I would like to acknowledge for their excellent contribution our collaborators in

Stockholm, Prof. L. Terenius and Dr. V. Vukojević, with the help of whom part of the data included in this thesis have been generated. The latter I especially thank for her belief in all that we can accomplish, her commitment to our project, her effort to interpret, assess and understand our findings, her never-ending passion for science, but most of all for her friendship and her patience.

Last but not least, I would like to thank Prof. Y.-A. Barde and Dr. V. Nikoletopoulou for their dedicated collaboration in the “retina” project (not included in this thesis), their expertise and motivation that allowed us to produce outstanding results using a technically unknown to us animal system.

All the above great scientists I thank from the bottom of my heart for allowing me to share, discuss, understand, enjoy and, finally, produce what one would justly call a tiny thread in the never ending “weave-world” of Biology.

Basel, 14th of June, 2010,
Dimitrios Papadopoulos

ABBREVIATIONS

aa: amino acids

Abd-A: *Abdominal A*

Abd-B: *Abdominal-B*

Antp: *Antennapedia*

ANT-C: *Antennapedia* Complex

AP: anterior-posterior

APD: Avalanche Photo Diodes

AP-1: Activator protein 1

ARFs: Auxin Response Factors

bcd: *bicoid*

Bip2: Bric-à-brac interacting protein 2

BX-C: *Bithorax* Complex

bp: base pairs

C: cysteine

C-terminus: carboxy-terminus

ceh-13: *C. elegans homeobox 13*

Dbl: Diffuse B-cell lymphoma

Dfd: *Deformed*

dlg: *discs large*

Dll: *Distal-less*

DNA: deoxyribonucleic acid

egl-5: *egg-laying defective 5*

En: Engrailed

ey: *eyeless*

Exd: Extradenticle

FCS: Fluorescence Correlation Spectroscopy

FLP: flippase

FoxP1: Forkhead box P1

Ftz: Fushi tarazu

GEF: Guanine nucleotide Exchange Factors

GFP: Green Fluorescent Protein

HD: Homeodomain

hh: *hedgehog*

HIV: Human Immunodeficiency Virus

HM: Homothorax-Meis

Hox: *Homeobox* (gene)

HOX-C: Homeotic Complex

hpo: *hippo*

Hth: Homothorax

Hsf: Heat shock factor

Hsp: Heat shock protein

hsp70: heat-shock protein 70

I: isoleucine

Lab: Labial

Lac repressor: Lactose repressor

lgl: *lethal giant larvae*

lin-39: *abnormal cell lineage 39*

M: methionine

mab5: *male abdominal 5*

Mad: Mothers against decapentaplegic

MAP: Mitogen-Activated protein

Mef2A: Myocyte enhancer factor 2A

N: asparagine

N: *Notch*

N-terminus: amino-terminus

NF-κB: Nuclear factor κB

nob-1: *knob-like posterior 1*

NtrC: Nitrogen regulatory protein C

N-WASP: Neuronal Wiskott-Aldrich Syndrome Protein

Oct-1: Octamer-1

Pax: Paired box

Pax6: *Paired box 6* (gene)

Q: glutamine

pb: *proboscipedia*

Pbx: Pre-B-cell leukemia transcription factor

PD: PAIRED domain

php-3: *posterior Hox gene paralogue 3*

Prep: Prolyl endopeptidase

R: arginine

Ras: *Rat sarcoma* (gene)

RepA: Replicative helicase A

RFP: Red Fluorescent Protein

RNA: ribonucleic acid

RNAi: RNA interference

RTK: Receptor Tyrosine Kinase

S: serine

sal: *spalt*

Sall1: *Spalt-like 1*

salm: *spalt major*

salr: *spalt related*

sav: *salvador*

Scr: *Sex combs reduced*

scrib: *scribbled*

Slp: Sloppy paired

Smad: Sma (*C.elegans*) mothers against decapentaplegic

Sp1: Stimulatory protein 1

SRF: Serum Response Factor

Tbp: TATA binding protein

T1: first thoracic segment (prothorax)

T2: second thoracic segment (mesothorax)

TALE: three amino acid loop extension

Tat: Trans-acting transcription factor

Tet: tetracycline (repressor or operator)

Tgf-β: Transforming growth factor β

TOR: Target of rapamycin

Tsc1: *Tuberous sclerosis 1*

Tsc2: *Tuberous sclerosis 2*

UAS: Upstream Activating Sequences

Ubx: Ultrabithorax

Vbp: Vitellogenin binding protein

Vnd: Ventral nervous system defective

W: tryptophan

w: *white*

wg: *wingless*

wts: *warts*

YPWM motif: tyrosine-proline-tryptophan-methionine motif

zen: *zerknüllt*

ABSTRACT

Over a century of research, invested in providing an answer to the fundamental question of how a single diploid cell constructs a whole organism, has placed *Drosophila* in the fore scene of developmental biology and has allowed us to merge the knowledge acquired from embryology with that of molecular biology and genetics.

The assignment of differential morphological and functional characteristics to the different body parts during development is largely mastered by *Homeobox* (*Hox*) genes which are involved in the specification of segmental identity along the anterior-posterior (AP) axis of bilateral animals. *Hox* genes encode transcription factors that contain the Homeodomain (HD), a helix-turn-helix DNA-binding domain, the conservation of which spans long evolutionary distances as the one between yeast and humans. The biological function of Hox transcription factors has been studied extensively to date and much is known about their role in development at the genetic, molecular and structural level. The molecular interactions of various HDs with their DNA binding sites have been precisely dissected *in vitro* and numerous developmentally important *Hox* downstream genes (effector genes) have been identified, providing the link between segmental specification and pattern formation, originally in *Drosophila* and subsequently in other animals.

However, the conservation of the HD in metazoans, albeit providing meaningful evidence for the interpretation of molecular evolutionary relationships between animal phyla, has been proven an obstacle in explaining the specificity of *Hox* target selection *in vivo*. Since all Hox orthologs and paralogs bind very similar regulatory sequences *in vivo*, it is challenging to explain how they simultaneously regulate differentially the formation of diverse body parts during development. The finding of a limited number of *Hox* cofactors has allowed us to take first steps towards solving this riddle, but a satisfactory explanation has not been provided yet.

In the present study we have constructed synthetic *Drosophila Hox* genes which encode a small carboxy-terminal (C-terminal) portion [tyrosine-proline-tryptophan-methionine (YPWM) motif, HD and C-terminus] of the full-length protein and have examined their function as transcriptional regulators genetically and their interactions with nuclear DNA in live cells, using methods with single-molecule sensitivity, quantitative imaging and Fluorescence Correlation Spectroscopy (FCS).

We have found that both *Antennapedia* (*Antp*) and *Sex combs reduced* (*Scr*) synthetic genes are functional *in vivo* and that the synthetic transcription factors find their binding sites primarily by multiple association/dissociation events, the rapidity of which is largely owed to electrostatic interactions.

Synthetic *Scr* genes exhibited specific function *in vivo* by inducing homeotic transformations in the embryo (allowing the formation of an additional pair of salivary glands) and in the adult fly (transforming the adult antenna into a prothoracic tarsus). They repressed antennal-specific genes, ectopically activated leg-specific genes in the antennal imaginal disc and bound DNA specifically both *in vitro* and *in vivo*. Their transformation capacity was found to be enhanced as compared to the full-length protein, suggesting that the amino-terminal (N-terminal) portion of the protein contributes quantitative, rather than qualitative, effects in Scr-mediated transcription.

Having proven their functionality *in vivo*, we used the synthetic *Scr* genes as tools for the study of HD-DNA interactions in live salivary gland cells, which express *Scr* normally during development, and thus represent a native Scr environment. By means of quantitative imaging, using Avalanche Photo-Diodes (APDs) with single-molecule sensitivity, we studied the molecular distribution and dynamics of synthetic transcription factors in polytene nuclei. We could distinguish wild type from mutant peptides directly by APD imaging, on the basis of their differential association with nuclear DNA, and study by FCS their movement and interactions with chromatin at the molecular level. This has been possible using the expression “leakage” of the *heat shock protein 70* (*hsp70*)-minimal promoter of the Upstream Activating Sequences (UAS)-constructs and its responsiveness to heat-shock, features, which facilitated measurements at low, physiologically-relevant levels. Thus, we were able to titrate the concentration of the transcription factor in live nuclei and to construct transcription factor-DNA binding curves in order to analyze the underlying chemical interactions kinetically. Using a simple two-step model in which non-specific interactions are followed by specific interactions, we derived experimentally the *in vivo* macroscopic equilibrium dissociation constant of the HD-DNA complex for specific and non-specific interactions, as well as the corresponding DNA-binding constants by numerical simulations.

Functional analysis was also performed using synthetic *Antp* genes. These were also able to confer gene-specific homeotic transformations during embryonic and larval development, such as head-to-trunk transformations reflected in the

embryonic cuticle, repression and activation of markers in the embryo, as well as antenna-to-mesothoracic-tarsus transformations in the adult, mediated by repression of antennal-specific genes and ectopic activation of leg-specific genes in the antennal imaginal disc. At the same time we examined the role of the Antp YPWM motif in transcriptional regulation and found it important for both activation and repression. We also investigated the importance of linker size between the YPWM motif and the HD (naturally varying in *Antp* splicing variants) and showed that linkers of different size preferentially favor or limit the function of the protein either as a transcriptional repressor or as an activator.

Taken together, our results show that synthetic *Hox* genes are functional *in vivo*. They exhibit gene-specific phenotypes, comparable to or enhanced over their full-length counterparts. This suggests that the long N-terminal portion of Hox transcription factors is not required for specificity, but for enhancing or limiting the strength of the transcriptional response. Our study also proposes synthetic genes as important tools for synthetic biology, since smaller but functional peptides might bear advantages for biomedically relevant applications, as compared to larger proteins. In addition, we establish an experimental platform for the quantitative study of transcription factor-DNA binding *in vivo*, devoid of overexpression and/or destructive observation of molecules and molecular interactions. We derive by FCS and quantitative imaging for the first time the *in vivo* HD-DNA binding constant and unravel the molecular interactions of the HD with chromatin in unprecedented detail. Our work might be used as a starting point for the dissection of transcription factor-DNA interactions *in vivo* which are relevant for development and disease.

INTRODUCTION

This section provides a summary of the achievements made in unraveling the molecular basis of development, emphasizing on *Hox* genes and their function in the specification of the animal body. At the same time, we discuss the appropriateness of *Drosophila melanogaster* as a model organism to study development and disease and the advent of molecular biology which enhanced our understanding about gene regulation in development. The essentiality of *Hox* genes for the determination of segmental identity is explained using examples of *Hox* gain- and loss-of-function mutations, which support the existence of a ground segmental state that becomes morphologically and functionally modified during evolution, as the number of *Hox* genes and their functional specificity increase vertically from lower to higher organisms. Moreover, we underscore the conservation of *Hox* genes among animal phyla using conclusions derived from the structural and functional dissection of the HD and uncover several considerations regarding the incomplete knowledge about transcriptional specificity of Hox transcription factors. As an answer to this problem, *Hox* cofactors are introduced and their function and conservation as transcriptional coregulators is outlined in detail, whereas the evolution of *Hox* gene complexes in animals is discussed from a molecular evolutionary perspective.

In a second step, we briefly give the background knowledge regarding the molecular interactions of the HD with DNA and explain the necessity of live cell experimental systems. Finally, the potential and appropriateness of FCS and the essentiality of tools for the quantification of macromolecular interactions in such studies are described, while certain achievements in the study of transcription factor-DNA interactions using FCS are presented. A century of fly genetic research and the availability of state-of-the-art methodologies allow nowadays the molecular dissection of transcriptional systems in their finest details and thereby the study of processes involved in development and disease.

THE GENETIC CONTROL OF DEVELOPMENT

One of the most remarkable wonders of life that has given rise to the interdisciplinary science of developmental biology has been the succession of events that lead to the formation of a whole organism from a fertilized egg. However, before the discovery of the DNA as the source of genetic information which can be inherited

from one organism to another, development could not be linked to genes.

The long lasting dispute between embryologists and geneticists in the first half of the 20th century has gradually led to the realization that genes contain all necessary information for morphogenetic processes that govern development, such as growth and differentiation. Although a solid confirmation of this belief came with the development of an adult animal upon transfer of somatic cell nuclei of the frog *Rana pipiens* into a nucleus-depleted zygote – performed by Robert Briggs and Thomas King (King and Briggs, 1956) – preliminary evidence was provided by the astute work of Thomas Morgan in the early 1900s. By isolating a mutant fly with white eyes and linking the gene responsible for this mutation (*white*; *w*) to the X chromosome (Morgan, 1910), he paved the road towards a revolutionized concept of heredity by linking its elements (genes) to chromosomes. Moreover, he and his student Hermann Muller showed that X-rays induce mutations and gave birth to genetics as a science that associates genes to phenotypes.

Certain conceptual advances thereafter enhanced our understanding about the genetic control of development. The pioneering work of E. Lewis on the Bithorax Complex (BX-C) and homeotic transformations (reviewed in (Lewis, 1994)) was based on initial studies of C. Bridges who used a collection of mutations that displayed a mild transformation of the third thoracic segment of the fly into a second thoracic segment, thus displaying a small second pair of wings [reviewed in (Lewis, 1994)]. These observations, acquired already before the molecular era of developmental genetics, allowed the gradual movement from gene structure to gene function (Lipshitz, 2004) and contributed to linking genes to development (Lewis, 1978). Moreover, the study of the correlation between gene function and pattern, initiated by C. Stern through the generation of genetic mosaics and extended by A. Garcia-Bellido and J. Merriam (Garcia-Bellido and Merriam, 1971) with the discovery of temporally controlled mosaics in imaginal discs, permitted the building of an important conceptual link between genes and cells and boosted our understanding about the organization of cells in space. Further work of E. Wieschaus and W. Gehring (Wieschaus and Gehring, 1976) tracked clonal lineages (mosaics) also earlier in development, in the embryo, and associated them to imaginal discs. Ever since, the flippase (FLP) system (Golic and Lindquist, 1989; Perrimon, 1998) and the use of histological markers in lineage analysis (Lawrence, 1981; Xu and Rubin, 1993) rendered mosaic analysis an indispensable tool of modern fly genetics (Arias,

2008).

However, it was not until the advent of molecular biology that important features of genes and their relation to development could be studied in detail. The molecular era in *Drosophila* genetics was initiated by the effort to obtain high resolution chromosomal maps of the BX-C and Antennapedia Complex (ANT-C) – reviewed in (Rubin and Lewis, 2000). Chromosomal walking and its usage in unraveling the molecular maps of the homeotic complexes (Bender et al., 1983a; Bender et al., 1983b; Garber et al., 1983), as well as the detection of their gene transcripts *in situ* (Akam, 1983; Hafen et al., 1983) helped associate transcriptional patterns to the genetic arrangement of their corresponding genes. At the same time, the discovery of transgenesis methods in the fly through P-element mediated transformation (Rubin and Spradling, 1982) and most importantly the development of the enhancer trapping methodology as a tool for identifying novel genes and analyzing their function (Bellen et al., 1989; Bellen et al., 1990; Grossniklaus et al., 1989; O'Kane and Gehring, 1987; Wilson et al., 1989) greatly increased the pace of genetic research and resulted in a large production of analytical tools, nowadays available to the broad scientific community.

Meanwhile, important contributions in the isolation of a large collection of mutants (Nusslein-Volhard and Wieschaus, 1980) shifted the attention of fly geneticists from the adult fly to the embryo where information about signaling, transcription, the cell cycle, the cytoskeleton and cell adhesion were reflected in the pattern of the embryonic cuticle (Arias, 2008) and allowed scientists to outline gene regulatory networks during early development (Ingham, 1988). Finally, the discovery of the conservation of the homeodomain in metazoans (McGinnis et al., 1984a; McGinnis et al., 1984b; McGinnis et al., 1984c; Scott and Weiner, 1984) initiated our understanding regarding the universality of genes and genomes and contributed immensely to the development of molecular evolution.

Although findings in genetics and development reach much farther from their immediate relevance to human biology, the evolutionary conservation of genes, gene networks and the mechanisms that these deploy to construct regulatory and signaling pathways in development has additionally resulted in the usage of diverse organisms to address questions regarding also homeostasis and disease in humans. *Drosophila* as a model organism has remained in the fore scene of such research, since it bears certain advantages over other animal models (to be discussed in the

following section).

Taken together, the advent of molecular biology, RNA and protein localization techniques, the technologies to delete the function of genes and to generate transgenic animals have been the critical findings in *Drosophila* research that have enhanced our current understanding about embryonic development, as a series of morphological changes that depends on the combinatorial action of differential gene expression, cell-cell communication and positional cues.

DROSOPHILA AS A MODEL ORGANISM FOR STUDYING DEVELOPMENT AND DISEASE

Its rapid life cycle, its large collection of mutants, the rapid genetic manipulation (Fig. 1) and the abundance of transgenesis techniques, including in the last few years specific gene targeting by homologous recombination (Matthews et al., 2005; Venken and Bellen, 2007), have rendered *Drosophila* a widely used experimental model system, originally for addressing questions of developmental interest, but also as a tool for studying cell biology, neuroscience and disease. Moreover, the completion of its genome sequencing (Adams et al., 2000; Myers et al., 2000) together with the addition of most heterochromatic sequences (Hoskins et al., 2007) have resulted in the biology of *Drosophila* being the most closely related invertebrate biology to humans (Boutros and Perrimon, 2000).

The notion that *Drosophila* could indeed be used as a system for understanding human genetics and subsequently the application of this knowledge for useful purposes originated from the finding that many genes involved in establishing body axes, cell types and organ systems have been highly conserved during animal evolution (Bier, 2005). Not surprisingly, such orthologous genes in the fly can functionally replace their vertebrate counterparts. In addition, flies have also been shown to offer a suitable experimental system for diverse studies, such as those of learning and behavior (Margulies et al., 2005), or even alcoholism and addiction (Bellen, 1998). Finally, approximately 700 human disease genes are estimated to have sufficiently well-conserved homologues to be analyzed in *Drosophila*. Such disease genes are categorized in five major groups: developmental defects (blindness, deafness, skeletal malformations, neural pathfinding disorders), neurodegenerative diseases, cancer, cardiac disease and immunological disorders – reviewed in (Bier, 2005). Here, a few important examples of human genetic

disorders, in which fly genetics has assisted in understanding their manifestation, are outlined.

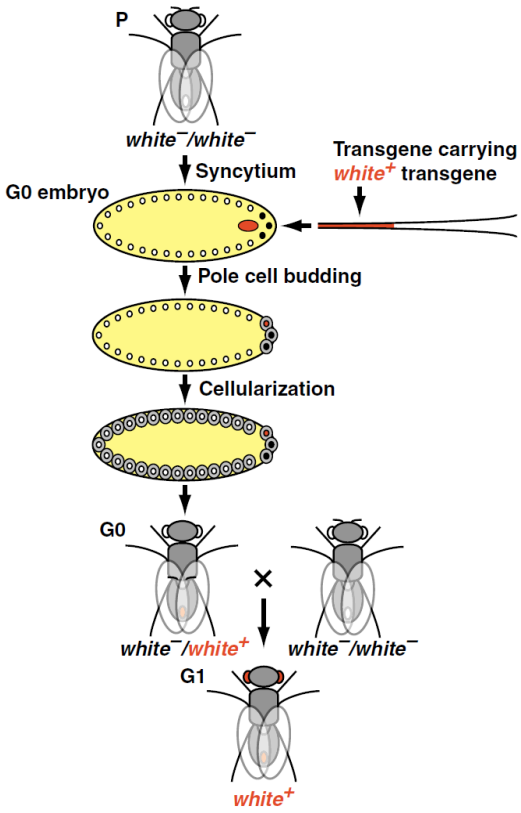

Fig. 1. *Drosophila* transgenesis. *white*$^+$ transgene DNA (red) is injected into generation zero *Drosophila* embryos (G0) of less than 1 hour old, which have been obtained from a parental (P) generation. The early developmental stages of *Drosophila* embryos are characterrized by rapid nuclear divisions that occur without accompanying cell divisions, creating a syncytium. Prior to cellularization, pole cells (black) bud off at the posterior end. For germ line transmission to occur, the transgenic DNA must be taken up into the pole cells that are fated to become germ cells. Transgenic DNA integrated into a pole cell (red pole cell) can be transmitted from one generation (G0) to the next (G1 progeny). The resulting integration events are identified using an appropriate marker, such as *white*$^+$. When used in a mutant *white*$^-$ strain, this transgene marks transgenic flies by giving them a darker eye color (Venken and Bellen, 2007).

Focusing on developmental disorders, these can be categorized in two major groups: the ones owed to genes that function in the same way in flies and humans and the ones that depend on genes that take part in a conserved developmental pathway, but control different processes in the two animals. Characteristic examples of the first group are the *Hox* genes that control segmental identity in both flies and human (the loss-of-function of which causes syndactyly and spinal cord defects in humans) – reviewed in (Carpenter, 2002) and references therein; or Paired box 6 (*Pax6*) [*eyeless* (*ey)* in the fly] (Czerny et al., 1999; Quiring et al., 1994) and *Spalt-like 1* (*Sall1*) (*spalt major* (*salm*) and *spalt related* (*salr*) in the fly) (Kohlhase, 2000) which cause eye development and auditory defects, respectively, in both organisms. To the second group belongs for example the *Notch* (*N*) receptor, which is necessary in vertebrates for the proper segmentation of the somatic mesoderm – reviewed in (Christ et al., 1998) – whereas in flies it plays an essential role in the patterning the drosophila wing and wing veins, as reviewed, for example, in (De Celis, 2003).

Drosophila can also be used as a model to study neurological disorders (Bodai et al., 2003; Bonini and Fortini, 2003; Driscoll and Gerstbrein, 2003; Muqit and Feany, 2002; Shulman et al., 2003). Second-site modifier screens have played an important role in the identification of genes that function in the same or overlapping pathways with the gene of interest (Fig. 2). In triplet-repeat diseases where the expansion of triplet repeats in relevant genes leads to neuronal degeneration, as in the case of Huntington disease (Rubinsztein, 2002), such screens can be used to identify genes that might play a role in reversing the degeneration phenotype. Proteins identified so far include the Heat shock (Hsp) proteins, which can act as chaperonins and assist other proteins in folding (Fig. 2). The same paradigm applies to Parkinson disease, where the, normally localized in the presynaptic terminals of dopaminergic neurons, α-synuclein protein aggregates in the cytoplasm (Maroteaux et al., 1988). Although no obvious homolog of α-synuclein is found in the fly, overexpression of the mutant human ortholog in the *Drosophila* eye disc results in neurodegeneration (Feany and Bender, 2000), a result which is partially reversed upon coexpression of the fly *park* gene (Haywood and Staveley, 2004; Yang et al., 2003), the fly ortholog of the human ubiquitin E3-ligase *parkin* (Shimura et al., 2000; Shimura et al., 2001). Further neurological disorders in which *Drosophila* has been used as a model is the familial Alzheimer disease (Shulman et al., 2003) and the fragile X syndrome (Ishizuka et al., 2002; Jin et al., 2003; Zhang et al., 2001).

Fig. 2. Examples of second-site modifier screens in *Drosophila melanogaster*. Suppression of polyglutamine induced retinal degeneration. A control eye expressing the human HSP70 protein (HSPA1L) under the control of the GMR driver (an enhancer that drives expression specifically in the eye) (left panel). A degenerating eye expressing a protein with an expanded polyglutamine domain MJDtr-Q78, under the control of GMR. This mutant protein contains a domain of 78 glutamines in place of the normal (Warrick et al., 1999) run of 27 residues (middle panel). Coexpression of the human chaperonin HSPA1L with MJDtr-Q78 results in suppression of the retinal degeneration phenotype caused by MJDtr-Q78 alone (right panel) (Warrick et al., 1999).

Finally, the fruitfly as a model system has contributed immensely to cancer research. Components of conserved signaling pathways today known to participate

in different types of human cancer, such as the Receptor Tyrosine Kinase (RTK) (Mulligan, 2001), *wingless* (*wg)* (Barth and Nelson, 2002), *hedgehog* (*hh*) (Bale, 2002), *Transforming growth factor β* (*Tgf-β*) (Siegel and Massague, 2003) and *N* pathways (Radtke et al., 2005) have functional homologs in the fly which play an important role in growth control decisions and/or the cell cycle. Meanwhile, overactivation of the insulin-branch of the Target Of Rapamycin (TOR) pathway caused by mutations in the *Tuberous sclerosis 1* (*Tsc1*) and *Tuberous sclerosis 2* (*Tsc2*) genes (Pan et al., 2004) also causes tumors in humans (Saucedo and Edgar, 2002). Also in this case, *Drosophila* has helped us gain insight into the molecular mechanisms controlling cell growth that are associated with this type of cancer. Furthermore, it has assisted in the identification of negative regulators of the cell cycle, such as *warts* (*wts*) (Justice et al., 1995; Xu et al., 1995), *salvador* (*sav*) (Tapon et al., 2002) and *hippo* (*hpo*) (Harvey et al., 2003), the functional homologs of which have been identified in mammals as important players in tumorigenesis. Subsequent analysis allowed the identification of the physical interaction between the corresponding proteins (Harvey et al., 2003). Finally, oncogenic behavior has additionally been observed in imaginal discs due to mutations in the tumor suppressor genes *lethal giant larvae* (*lgl*), *scribbled* (*scrib*) and *discs large* (*dlg*) or in an activated oncogenic form of *Rat sarcoma* (*Ras*) gene. However, the combination of both mutations in mitotic clones results in epithelial-to-mesenchyme transitions and metastatic tumors throughout the fly body (Brumby and Richardson, 2003; Pagliarini and Xu, 2003). The relevance of these genes and their mutations to vertebrate growth control and proliferation has also been described (Dow et al., 2003; Skouloudaki et al., 2009).

Taken together, the above studies prove *Drosophila* a valuable genetic tool not only for understanding the principles of genetics in animals, but also as an easy-to-manipulate experimental system that facilitates biomedically relevant studies, some of which would not be possible to perform in mammalian model systems, or at least not as efficiently or rapidly as in the fly.

One obvious question stemming from these observations is whether *Drosophila* could provide answers to any question regarding human biology. The answer is obviously negative, since there are physiological or developmental processes in vertebrates which have been “invented” later during animal evolution (e.g. adaptive immunity or the formation of the vertebrate limb, respectively), on

which information can be hardly obtained from the fly. Nevertheless, even in these cases, the fly can be and is used comparatively for understanding the evolutionary frame in which certain innovations are placed, as well as the adaptive advantages that allowed their establishment in higher organisms.

HOX GENES AND THE DETERMINATION OF SEGMENTAL IDENTITY IN ANIMALS

Among the first groups of fly genes found to share remarkable evolutionary conservation with orthologous genes found in both higher and lower animal phyla have been *Hox* genes, originally analyzed by E. Lewis as being the driving force of homeotic mutations in the fly (Lewis, 1978) and studied by many scientists thereafter. Even though *Hox* genes mark several decades of fly research as compared to other developmentally important genes, their mode of action as transcription factors is believed to be very complex, such that our knowledge regarding their developmental meaning remains incomplete to date.

Hox genes are essential for conferring segmental identity along the AP axis of bilateral animals (McGinnis and Krumlauf, 1992). They are expressed in a linear manner along the different segments of the animal body, in the same way they are arranged one after the other along the chromosome (Fig. 3).

This rule or pattern of organization was first described by the term “colinearity” or “spatial colinearity” by E. Lewis and appears to be a general principle of *Hox* gene clusters, identified originally in *Drosophila* (Kaufman et al., 1980; Lewis, 1978) and later on in animals as diverse as worms (Kmita et al., 1998) and mammals (Boncinelli et al., 1989; Duboule and Dolle, 1989). Remarkable as this level of organization may seem, it does not explain which functional constrains have requested that *Hox* genes are aligned in an ordered fashion. In parallel, except of “spatial colinearity”, the term “temporal colinearity” has been proposed to explain the fact that *Hox* genes are also expressed in a temporal succession during development (Kmita and Duboule, 2003) (at least in the mouse studied so far). The evolutionary persistence of *Hox* genes is reflected in the fact that ever since their identification in the fly more than 1000 *Hox* genes have been found in metazoans (Abzhanov and Kaufman, 2000; Cook et al., 2004; Levine et al., 1984; Manuel et al., 2006; Seimiya et al., 1994) and virtually all bilateral animals studied so far (Duboule and Morata, 1994), but also in other kingdoms like fungi and plants (Shepherd et al.,

1984; Sommer et al., 1990).

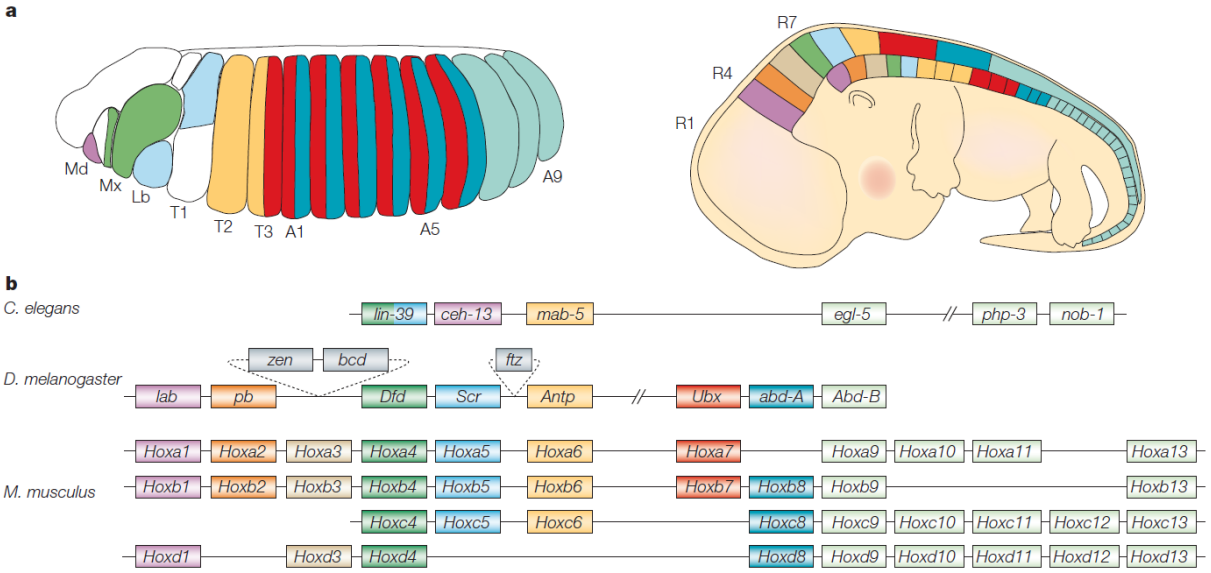

Fig. 3. Hox expression, genomic organization, and Hox binding sequences. (a) The panel on the left shows a stage 13 *Drosophila melanogaster* embryo that has been coloured in the schematic to indicate the approximate domains of transcription expression for all Hox genes except *proboscipedia* (*pb*) (Kosman et al., 2004). The segments are labelled (Md, mandibular; Mx, maxillary; Lb, labial; T1–T3, thoracic segments; A1–A9, abdominal segments). The panel on the right shows a mouse (*Mus musculus*) embryo, at embryonic day 12.5, with approximate Hox expression domains depicted on the head–tail axis of the embryo. The positions of hindbrain rhombomeres R1, R4 and R7 are labelled. In both diagrams the colours that denote the expression patterns of the Hox transcripts are colour-coded to the genes in the Hox cluster diagrams shown in (b). Anterior is to the left, dorsal is at the top. (b) A schematic of the Hox gene clusters (not to scale) in the genomes of *Caenorhabditis elegans*, *D. melanogaster* and *M. musculus*. Genes are coloured to differentiate between Hox family members, and genes that are orthologous between clusters and species are labelled in the same colour. In some cases, orthologous relationships are not clear (for example, *abnormal cell lineage 39* (*lin-39*) in *C. elegans)*. Genes are shown in the order in which they are found on the chromosomes but, for clarity, some non-Hox genes that are located within the clusters of nematode and fly genomes have been excluded. The positions of three non-Hox homeodomain genes, *zerknüllt* (*zen*), *bicoid* (*bcd*) and *fushi tarazu* (*ftz*), are shown in the fly Hox cluster (grey boxes) (Pearson et al., 2005).

Drosophila features a single *Hox* gene cluster which has been split in two complexes, the ANT-C) and the BX-C (Kaufman et al., 1980), but vertebrate *Hox* genes have been duplicated twice early in vertebrate evolution (Holland et al., 1994) and are comprised of clusters of nine to eleven orthologs. Most of the information available for the biological function of *Hox* genes comes from gain-of-function and loss-of-function experiments that transform one body segment or part into another. Such homeotic transformations are described extensively in (Gehring et al., 2009). All loss-of-function mutations in the *Drosophila* Homeotic Complex (HOX-C) lead towards the formation of a mesothoracic segment, whereas *Hox* gain-of-function transforms the rest of the body segments into a non-mesothoracic fate. This principle is both reflected in the developing embryo as readout of the different morphological

characteristics of each segment and in the larva and adult fly as transformations of imaginal discs (and thus body parts) into mesothoracic discs, that normally give rise to a pair of wings and a pair of second thoracic segment (T2) legs (Fig. 4) (Gehring et al., 2009).

These phenotypic observations are important in that they both provide experimentally-derived evidence about the existence of a ground state in segment specification of bilateral animals and explain the diversity of segments observed in dipteran insects, as compared to other more primitive insects, arthropods or even lower animal phyla, such as annelids, where little (if any) diversification of the different body segments is observed. This ground segmental state corresponds to T2 (mesothoracic segment), originally proposed by E. Lewis for the BX-C and experimentally supported by W. Gehring for the entire HOX-C (Gehring et al., 2009). It bears a pair of wings (dorsal appendages) and a pair of middle legs (ventral appendages). Crustaceans, to which insects are most closely related (Friedrich and Tautz, 1995; Giribet et al., 2001; Shultz and Regier, 2000), display a pair of legs throughout their abdominal segments, the formation of which is thought to be suppressed in higher arthropods like insects by means of posterior to T2 *Hox* genes. In an analogous way, the function of prothoracic *Hox* genes normally expressed anterior to the T2 segment accounts for the phenotypic suppression of the prothoracic wings, found in fossils of more primitive insect species, both in prothoracic and abdominal segments (Carroll et al., 1995; Gehring et al., 2009).

Hox gene	Ortholog	Phenotype	Transformation
Labial (*lab*)	(Hox1)	Defects in all gnathocephalic segments. Expression in the intercalary or most anterior gnathal segments	Head defects
Proboscipedia (*pb*)	(Hox3)	Labial palps transformed to tarsal structures (T1)	Posterior ↓
Deformed (*Dfd*)	(Hox4)	Defects in mandibular and maxillary segments. Maxillary sense organ mouth hooks, maxillary cirri (missing). Head → thorax	Head defects posterior ↓
Sex combs reduced (*Scr*)	(Hox5)	T1 → T2	Posterior ↓
		Mx ← Lb	Anterior ↑
Antennapedia (*Antp*)	(Hox6)	T2 → T1 (PS4 and 5)	Anterior ↑
Ultrabithorax (*Ubx*)	(Hox7)	A1, T3 → T2	Anterior ↑
abdominal-A (*abd-A*)	(Hox8)	A2-A8 → A1	Anterior ↑
Abdominal-B (*Abd-B*)	(Hox9)	A5, A6, A7 → anterior transformation	Anterior ↑

Hox gene	Ortholog	Phenotype	Transformation
Labial (*lab*)	(Hox1)	Head defects, missing abdominal segments	Head defects
Proboscipedia (*pb*)	(Hox3)	Additional antennomaxillary sense organs on thoracic and abdominal segments	Anterior ↑
Deformed (*Dfd*)	(Hox4)	Additional maxillary cirri on T1 and T2	Anterior ↑
Sex combs reduced (*Scr*)	(Hox5)	T2, T3 → T1	Anterior ↑
Antennapedia (*Antp*)	(Hox6)	T1 → T2	Posterior ↓
Ultrabithorax (*Ubx*)	(Hox7)	H, T1, T2, T3 → A1	Posterior ↓
abdominal-A (*abd-A*)	(Hox8)	T1, T2, T3 → A	Posterior ↓
Abdominal-B (*Abd-Br*)	(Hox9)	T1-A7 → A8/9	Posterior ↓
		T1-A7 → A8/9	
(*Abd-Bm*)		T1-A8 → A9 (Filzkörper)	Posterior ↓

Fig. 4. Epistatic relationships between different *Hox* paralogs. Loss-of-function mutations lead to transformation of segments anterior or posterior to T2 into T2 (left panel), whereas *Hox* gain-of-function superimpose information onto the default T2 developmental program, thus taking segmental identity away from T2 (right panel) (Gehring et al., 2009).

In fact, expression of *Scr* in the prothorax is believed to have been responsible in higher insects to phenotypically suppress the formation of wings and there is certain amount of evidence that supports this hypothesis (Beeman et al., 1989; Carroll et al., 1995; Chesebro et al., 2009; Rogers et al., 1997). Recent experiments in which the function of *Scr* was depleted by RNA interference (RNAi) in *Oncopeltus fasciatus*, the large milkweed bug (Chesebro et al., 2009), and in the cockroach *Periplaneta americana* (Hrycaj et al., 2010), have demonstrated the formation of ectopic prothoracic wings which "deletes" the information imposed by segment modification genes (in this case *Scr*) and allows the recapitulation of the ancestral prothoracic segment. These studies have increased significance, since they have been performed on hemimetabolous insects (more primitive than *Drosophila*), a result suggesting that the adaptive suppression of first thoracic segment (T1) wings might have occurred inside the insect class.

The aforementioned homeotic mutations in the fly, where the various *Hox* genes act to modify the default T2 identity, suggest the developmental basis of the ground state, whereas the gradual transformation in the course of evolution of anterior and posterior segments away from a T2 identity argues in favor of an evolutionary ground state (Gehring et al., 2009). Thus, it is tempting to speculate that the developmental and evolutionary ground states are in line with Ernst Haeckel's belief in 1874 that "ontogeny recapitulates phylogeny", a declaration that has faced both scientific and social criticism ever since, but has began gaining support after the development of molecular biology and the realization that evolution acts by modifying existing patterns, rather than by innovation (Gehring et al., 2009) – reviewed in (Johnston et al., 1992).

THE HOMEODOMAIN: CONSERVATION, STRUCTURE AND SPECIFICITY

A 180 base pairs (bp) long DNA sequence present in the coding region of *Hox* genes was found similar to other fly genes known to play an important role in *Drosophila* development (McGinnis et al., 1984c; Scott and Weiner, 1984) and was termed the homeobox. It encodes a conserved 60 amino acid (aa) domain, the HD, nowadays known to be the conserved DNA-binding domain of Hox proteins that evolutionarily spans more than the whole animal kingdom, ranging from yeast to humans – reviewed in (Gehring, 1987). To date, several three-dimensional

structures of DNA-bound HDs of various HD-containing proteins have been resolved. These include Antp (Qian et al., 1989), Engrailed (En) (Kissinger et al., 1990), Ftz (Qian et al., 1994a), Scr (Joshi et al., 2007), Ultrabithorax (Ubx) (Passner et al., 1999; Rohs et al., 2009) and HoxB1 (Piper et al., 1999). HDs that share 60% or more similarity with the Antp HD are traditionally classified as the Antp-type of HDs; they display identical sequences in helix III of the HD and the conserved YPWM motif N-terminally to the HD (Akam, 1989; Krumlauf et al., 1986; Mavilio et al., 1986; Regulski et al., 1985). Occasionally, structural similarities are also present in other parts of Hox orthologs, as for example in Deformed (Dfd) (Graham et al., 1988; Regulski et al., 1987).

The structure of the Antp HD showed that it comprises a helix-turn-helix DNA binding domain, consisting of four α-helices and an N-terminal flexible arm. Its helix-turn-helix motif is very similar to prokaryotic DNA-binding transcription factors, the helix-turn-helix motif of which also binds DNA in a sequence-specific manner (Laughon and Scott, 1984; Otting et al., 1988; Shepherd et al., 1984). The third helix (also termed the recognition helix) forms base-specific contacts with the major groove of the DNA, the N-terminal arm interacts sequence-specifically with the DNA minor groove and the turn between helix I and helix II with the DNA-backbone (Otting et al., 1988), as demonstrated in helix-turn-helix swapping experiments in Ftz (Furukubo-Tokunaga et al., 1992). Specific contacts of the recognition helix with the DNA major groove are mediated by three residues, isoleucine (I) 47, glutamine (Q) 50, asparagine (N) 51 and methionine (M) 54 (Qian et al., 1989). Moreover, the hydrophobic core of the HD is conserved in most of the HDs analyzed (Kissinger et al., 1990; Scott et al., 1989). A structural representation of the Antp HD-DNA complex is presented in Fig. 5.

The specificity of the HD of at least Antp and Scr is attributed to the flexible N-terminal arm, which contains two highly conserved arginine (R) residues at positions 3 and 5 (Joshi et al., 2007; Qian et al., 1989). In fact, the specificities of Antp and Scr in the fly are owed almost exclusively to their corresponding HDs (Gibson et al., 1990). This result does not come as a surprise if we consider that the HDs of these two Hox proteins differ in only five residues, four of which reside in their N-terminal arms, and the interchange of these residues changes the specificities of the two proteins *in vivo* (Furukubo-Tokunaga et al., 1993; Zeng et al., 1993). The specificity of the N-terminal arm of the HD seems to be a general feature of Hox proteins, since

the same principle applies to the N-terminal arms of Dfd and Ubx (Chan and Mann, 1993; Lin and McGinnis, 1992; Mann and Hogness, 1990). Also, evidence supporting this notion stems from the observation that residues 6 and 7 of the Scr HD are selectively phosphorylated, rendering the HD unable to bind DNA specifically both *in vivo* and *in vitro* and to exhibit gene-specific phenotypes in the fly (Berry and Gehring, 2000; Papadopoulos et al., 2010; Vukojevic et al., 2010). The latter two studies demonstrate that the functional specificity of Scr is limited to a small C-terminal portion of the protein, containing the HD and the YPWM motif, which has been termed the synthetic Scr peptide. Recent findings have extended this knowledge to synthetic Antp peptides, which encode an analogous portion of the protein (present study).

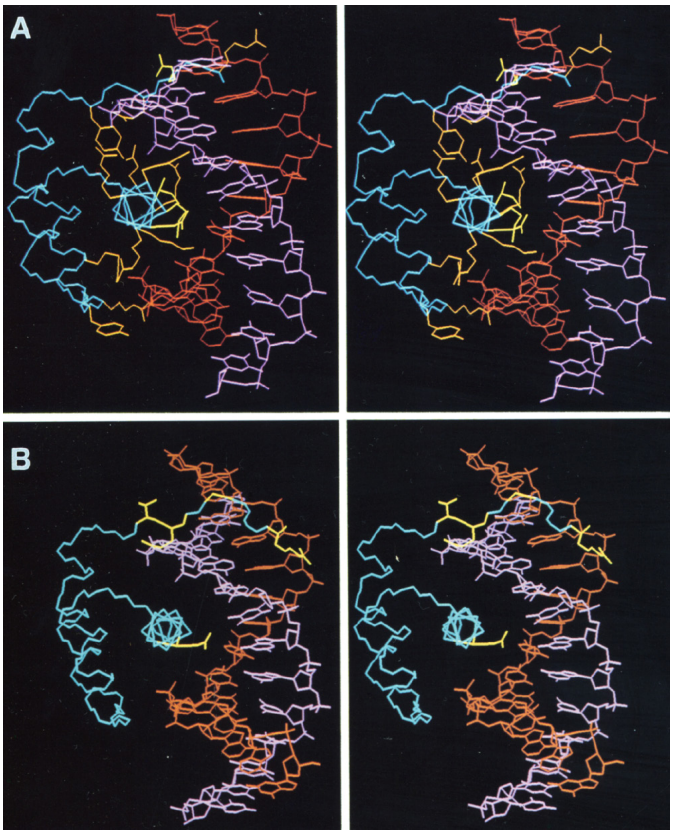

Fig. 5. Stereo views of one of the 16 conformers used to describe the NMR solution structure of the Antp(C39S) homeodomain-DNA complex. (A) The drawing shows the backbone of the homeodomain residues 3-55 (cyan), all side chains of the homeodomain that contact the DNA (yellow; for the selection criteria used for inclusion of individual side chains, see (Billeter and Wuthrich, 1993)), and the base pairs 3-13 of the DNA (red for the α strand with the sequence d-GAAAGCCATTAGAG, magenta for the complementary β strand). (B) Stereo view of one conformer, emphasizing those aa residues (shown in yellow) that are implicated in the DNA binding specificity (Q 50) and in the functional specificity (R 1, G 4, T 7) by genetic experiments (Gehring et al., 1994).

The sequence analysis of the Antp HD has shown that nine aa are invariant among many HDs examined, ranging from sea urchin to human, whereas the rest of the aa are most closely related to the Antp HD aa sequence (Gehring, 1987), a finding which supports the notion that homeoboxes found in *Hox* genes that reside in the middle of the HOX-C (like *Antp*) are most closely related to the ancestral HD sequence, whereas HDs found in more anterior or posterior *Hox* genes are less similar in sequence to the ancestral gene (Gehring et al., 2009). Interestingly, bacterial or

fungal HDs which share noteworthy structural similarities with HDs of higher organisms are contained in proteins that also control processes of cellular differentiation (Schulz et al., 1990; Shepherd et al., 1984).

REQUIREMENT OF HOX COFACTORS FOR INCREASED SPECIFICITY *IN VIVO*

Although all HDs bind DNA sequence-specifically *in vitro*, the need for increased specificity for the *in vivo* regulation of their targets soon became evident in *Hox* research, owing to the fact that most Hox proteins bind simple sequences with similar affinities *in vitro* (Ekker et al., 1994; Hoey and Levine, 1988). The paradox of Hox specificity could be summarized in two different questions: “Since Hox proteins bind very short and similar sequences in the regulatory regions of genes and these sequences are statistically encountered very frequently in the genome (i.e. several fold more frequently than the plausible maximum amount of genes that they could potentially regulate), what determines the selection of genes that will be regulated (repressed or activated) by a given Hox transcription factor at a given time and tissue during development?”; and “Which is the decisive factor that determines which Hox paralog will regulate a given gene, if they all bind so similar sequences *in vivo*?”. These questions have not been satisfactorily answered to date, but several important steps have been taken towards providing a relevant mechanism of action. Since there could not be possible reasons why one would doubt the DNA-binding preferences for Hox proteins obtained by *in vitro* and structural studies, other factors should account for increasing the specificity of Hox proteins *in vivo*. Extradenticle (Exd), a Three Amino acid Loop Extension (TALE) HD-containing protein was first identified by genetic studies to be an actual cofactor of Hox proteins (Mann and Chan, 1996), since mutations in the *Exd* gene produced homeotic phenotypes, without affecting *Hox* expression patterns (Peifer and Wieschaus, 1990). *Exd* encodes a member of the Pre-B-cell leukemia (Pbx) family of proteins and was found to be conserved from *C. elegans* to mammals (Bodai et al., 2003; Mann and Chan, 1996; Moens and Selleri, 2006; Van Auken et al., 2002).

Thereafter, a second cofactor, Homothorax (Hth), also belonging to the TALE group of HD proteins, was found necessary for the nuclear localization of Exd (Berthelsen et al., 1999; Kurant et al., 1998; Rieckhof et al., 1997) and essential to form a heterotrimer with Exd and Hox, in which all transcription factors are bound to

the DNA and protein-protein interactions occur between Hox and Exd, as well as between Exd and Hth (Akin and Nazarali, 2005; Mann and Affolter, 1998; Mann and Morata, 2000; Moens and Selleri, 2006), increasing thus the specificity of target selection. Functional dissection of the Hth-Exd complex showed that their interaction is DNA-dependent and forms between the PBC domain of Exd (Pbx) and the Homothorax-Meis (HM) domain of Hth (Mann and Affolter, 1998). The vertebrate orthologs of Hth belong to the Meis and Propyl endopeptidase (Prep) classes of proteins (Berthelsen et al., 1998; Burglin, 1997) and Meis proteins have been found to interact also directly with Hox peptides, suggesting that they also behave as putative Hox cofactors (Shen et al., 1997a; Williams et al., 2005).

Ever since, physical interactions between Hox transcription factors and other proteins that either assist or prevent the former from triggering a transcriptional response, have been described. To the first group belongs Bric-à-brac interacting protein 2 (Bip2) (Prince et al., 2008), which binds the Antp YPWM motif *in vivo* and allows it to interact with the basal transcriptional machinery and in the second category fall Paired box (Pax) transcription factors like Ey in the fly, the PAIRED Domain (PD) of which binds to the HD of Hox factors and prevents them from functioning as transcriptional regulators (Plaza et al., 2008; Plaza et al., 2001).

More transcription factors (also termed Hox collaborators) have been found to bind Hox binding sites in collaboration with Hox proteins, although no cooperative binding between them and the Hox peptides has been directly demonstrated. These include Sloppy paired (Slp) (Gebelein et al., 2004), collaborating with Ubx and Abdominal A (Abd-A) to repress *Distal-less* (*Dll*) in the abdomen; Forkhead box P1 (FoxP1) (a vertebrate ortholog of Slp), necessary for the establishment of motor neuron identities in the mouse (Dasen et al., 2008; Rousso et al., 2008); Sma mothers against decapentaplegic (Smad) proteins (Galant et al., 2002; Grienenberger et al., 2003; Shi et al., 2001; Shi et al., 1999; Walsh and Carroll, 2007); and finally Mothers against decapentaplegic (Mad)/Medea as collaborators of Ubx on the *spalt* (*sal*) *1.1* element, where they act as transcriptional repressors, or as collaborators of Abd-A to activate transcription of the *wg* gene (Grienenberger et al., 2003; Walsh and Carroll, 2007). Physical interactions between Hox factors and the aforementioned assistant proteins have not been described to date. The best described case of interactions includes those of the Pbx-Hox complex which are therefore discussed in more detail hereafter.

The binding of Pbx to Hox proteins is attributed to the physical interaction between the TALE motif of Pbx proteins and the YPWM motif of Hox proteins, in which the three aa of the TALE domain create a hydrophobic pocket which binds the tryptophan (W) of the YPWM, as supported by biochemical, structural and *in vivo* studies (Chan and Mann, 1996; Chang et al., 1995; Joshi et al., 2007; LaRonde-LeBlanc and Wolberger, 2003; Lu and Kamps, 1996; Neuteboom et al., 1995; Passner et al., 1999; Phelan et al., 1995; Piper et al., 1999). Even in Hox proteins that exhibit sequence variation in their YPWM motif, as in the case of Abdominal B (Abd-B), the W at position 3 is still conserved and participates in this interaction (Shen et al., 1997a), a piece of evidence that further explains the importance of this conserved residue. However, it might be the case that additional residues and/or Hox/Pbx protein domains stabilize the interaction between homeoproteins and cofactors, since mutation of the YPWM motif failed to abolish the TALE-YPWM interaction of Ubx and also part of Ubx function *in vivo* (Galant et al., 2002; Merabet et al., 2003; Shen et al., 1997a).

It is important to note that Hox-independent functions of *Exd* and Hth have also been identified, suggesting that TALE HD proteins participate in processes that are not linked to their function as Hox-cofactors (Bessa et al., 2008; Casares and Mann, 1998; Jiang et al., 2008; Laurent et al., 2007; Moens and Selleri, 2006). Inversely, several functions of Hox transcription factors *in vivo* are delivered in a cofactor-independent way (Peifer and Wieschaus, 1990; Rawat et al., 2008a; Shen et al., 2004a).

Last but not least, important conclusions can be drawn in conjunction with the requirement for cofactor-mediated *Hox* function *in vivo*. Comparison of 66 different *cis*-regulatory elements, for which there is substantial evidence for direct regulation by Hox factors *in vivo* or *ex vivo* – reviewed in (Mann et al., 2009) – revealed that an obvious Pbx input is present in genes activated, rather than repressed by Hox proteins. Second, a requirement for Pbx input is more often observed when the target gene is regulated by an anterior *Hox* gene (paralogs 1-5), rather than by a posterior one. Moreover, elements lacking obvious Pbx input tend to contain multiple *Hox* binding sites, a fact which may reflect the need of *Hox* binding sites to cluster in groups to increase the affinity of Hox proteins for regulatory elements when cofactors do not participate in the regulation of the gene in question. Finally, it is worth mentioning that the selectivity of Hox peptides themselves for a specific DNA binding

site sequence has been shown to be altered upon cooperative binding with a Pbx factor, as demonstrated for Labial (Lab) (Hox1) and Ubx (Berger et al., 2008; Noyes et al., 2008).

THE EVOLUTION OF THE *HOX* GENE COMPLEX IN METAZOANS

The epistatic relations among *Hox* genes in the modification of the ground segmental state, the conservation of the structure and specificity of their HD, as well as the universality of their mechanisms of action (e.g. cofactors) all suggest common ancestry in the molecular evolution of *Hox* genes. The hypothesis that *Hox* genes evolved from an ancestral HD-containing gene (the so-called "*urhox*" gene) is supported by the fact that *Hox* genes are fewer in evolutionarily lower organisms (such as cnidarian species (Finnerty and Martindale, 1997; Finnerty and Martindale, 1999; Finnerty et al., 2004; Gauchat et al., 2000; Ryan et al., 2007)) and they start to increase in number after the evolution of bilateral symmetry, when the acquisition of two distinct body ends (one anterior and one posterior) required the elongation of the body along an axis and the diversification of intermediate body parts.

Already cnidarians, such as *Nematostella vectensis*, which exhibit perhaps the most primitive form of bilateral symmetry, are thought to be intermediate organisms between radially symmetric and bilateral animals and display only a few Hox genes which are anterior or posterior, while lacking intermediate gene members (Finnerty et al., 2004; Gauchat et al., 2000). It is commonly accepted that the transition of radial to bilateral symmetry was driven by the need of predator organisms for directed locomotion, although alternative theories are also evolutionarily plausible (improvement of the efficiency of internal circulation by affecting the compartmentalization of the gut and the location of major ciliary tracts) – reviewed in (Finnerty, 2005). Nevertheless, originally, segmentation created repetition of identical body parts (as in the case of annelids), whereas higher animals (featuring more *Hox* genes) modified these segments either mostly internally (millipedes), or internally and externally (arthropods and higher animals). Therefore, it is conceivable that the addition of the intermediate *Hox* gene members is an event (or a series of events) relatively more evolutionarily recent in animal development.

A common mechanism for the generation of gene families of paralogous genes is gene duplication, but a mechanism by which genes are inserted in the

middle of the cluster must have been utilized in the evolution of *Hox* genes to explain the presence of only anterior and posterior genes in more primitive animals and the presence of middle *Hox* genes, in addition, in higher animals. One plausible scenario is gene duplication by unequal cross over (Gehring, 1998; Zhang and Nei, 1996), facilitated by repetitive DNA sequences flanking the "*urhox*" gene (Goldberg et al., 1983). This mechanism – reviewed in (Gehring et al., 2009) – would have two obvious features: it would allow that the intermediate *Hox* genes are the most similar in sequence to the *urhox* gene and that the ones at the very beginning or end of the cluster are the least similar to it, since their aa sequence is the least recombinant and has had the longest time to diverge from the ancestral sequences (Gehring et al., 2009). Fig. 6 outlines the *Hox* genes' phylogenetic network that is consistent with this evolutionary hypothesis.

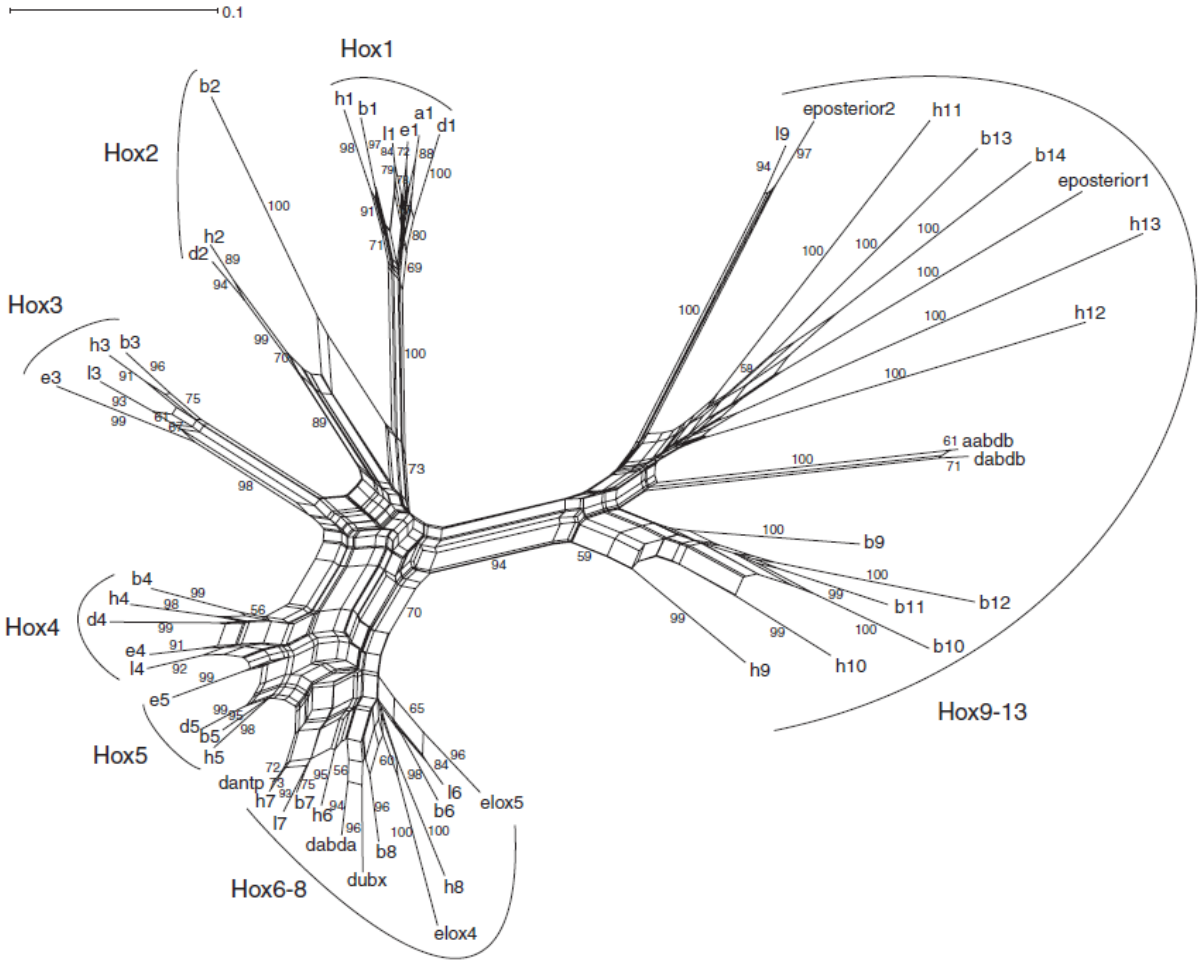

Fig. 6. **Phylogenetic network of the *Hox* cluster genes using the split-tree program (Huson and Bryant, 2006).** The anterior (*Hox1-8*) and posterior *Hox* genes (*Hox9-13*) are clearly separated. Whereas the *Hox1-5* genes are clearly separated, the intermediate genes *Hox6-8* are not resolved. This is in line with the assumption that intermediate *Hox* genes have arisen more recently in evolution and, therefore, have diverged the least (Gehring et al., 2009).

Three important conclusions can be derived regarding the phylogenetic relations between different *Hox* genes from different organisms: first, only orthologs 1, 2, 3, 4 and 5 can be clearly distinguished, whereas orthologs 6, 7 and 8 and, even more, 9-13 are poorly separated. Second, anterior *Hox* genes (paralog groups 1-8) share less sequence similarities with, and are therefore distant from, posterior *Hox* genes (paralogs 9-13), evidence which strongly supports the gradual insertion of *Hox* paralogs in the middle of the cluster during evolution as well as the increasing diversification of *Hox* members in both directions (towards the anterior and the posterior end of the complex). Finally, as the genomes of more organisms are being sequenced, one might be able to precisely determine the time points of the gene duplication events in *Hox* evolution that have led to the formation of the *Hox* gene complexes of mammals and human, as we know them today.

SEQUENCE-SPECIFIC HOMEODOMAIN-DNA BINDING *IN VITRO*

Although the structural studies of several HD-DNA complexes of various homeoproteins provide important evidence about the three-dimensional conformation of the HD upon DNA-binding, the residues playing an important role in HD-DNA recognition and the binding-sequence preferences of different HDs, they are not suitable for describing the dynamics of the binding process, viewed as a reversible chemical reaction, which contains information about the duration of the transcriptional response and the relevant concentrations in which transcription factors exhibit binding to their target sequences.

The first steps towards describing the dynamic binding process of the HD to DNA were taken using the purified Antp HD in solution and the putative Ftz-binding site BS2, which is normally situated in the *en* promoter, identified as a Ftz-protected site in footprinting experiments (Muller et al., 1988). This sequence represented a true Ftz-binding site, as shown in genetic experiments (Howard and Ingham, 1986) and since the Antp and Ftz HDs have identical recognition helix sequences, this binding site was bound also by the Antp HD. *In vitro* analysis by means of gel-retardation assays using a 68 aa Antp HD derived the first HD-DNA equilibrium dissociation constant for specific binding ($K_{D,specific}$) to be 12 nM (Muller et al., 1988). Subsequent analyses with an Antp HD, in which cysteine (C) 39 had been mutated to serine (S) to abolish artificial dimerization *in vitro*, corrected the initial dissociation constant value to 1.6 nM, while non-specific interactions were also determined in the

same study to lie in the range of 0.06-0.1 μM using a poly-cytosine oligonucleotide (Affolter et al., 1990).

These measurements, although differing approximately one order of magnitude from each other, are in agreement with subsequent studies which scored dissociation constants of 2.9-5.7 nM for specific interactions and 0.08-0.17 μM for non-specific interactions, using surface plasmon resonance measurements with the same original Antp peptide (Seimiya and Kurosawa, 1996). Moreover, they are in line with the binding properties of other HDs, such as the Hoxd1 HD (the ortholog of the *Drosophila* Lab), which exhibited a dissociation constant of 8.6 nM (Kumar and Nazarali, 2001); the *Drosophila* Ventral nervous system defective (Vnd) HD with a constant of 19 nM (Wang et al., 2002) and the Octamer-1 (Oct-1) HD with 16.3 nM (Doucleff and Clore, 2008). Taken together, this dataset suggests that various HDs bind with similar affinities to their binding sites *in vitro*, but at the same time these studies might have to be interpreted with caution, since they do not reflect the physiological conditions within a live cell, where a plethora of other contributions (positive or negative) are reflected in the macroscopic binding constants of transcription factors. A few of these contributions could be owed to chromatin structure and dynamics (histone modifications and DNA methylation), accessibility of the binding sites, presence or absence of transcriptional co-regulators, transcription factors competing with one another for binding and also physiological parameters, such as local pH and temperature in the proximity of the binding site. Since all these factors cannot be easily taken into consideration or reconstituted in solution, live-cell experimentation is indispensable for describing such multi-factorial processes as transcription factor-DNA binding and the initiation of a transcriptional response.

FLUORESCENCE CORRELATION SPECTROSCOPY AS A METHOD FOR STUDYING MACROMOLECULAR INTERACTIONS IN LIVE CELLS

One of the most suitable methods for investigating such dynamic processes with the minimum possible disturbance and very high sensitivity (reaching the detection capacity of single molecules) is FCS, a biophysical technique used for describing molecular interactions *in vitro* and *in vivo*. The method is based on recording and correlating fluorescence intensity fluctuations of single fluorescing molecules and extracting information regarding their mobility properties (molecular

transport and diffusion), molecular interactions and kinetics, as well as photophysical reactions occurring in the sample being analyzed.

Although the theoretical knowledge of FCS has been established more than 35 years ago (Ehrenberg and Rigler, 1974; Elson and Magde, 1974; Koppel, 1974; Koppel et al., 1976; Magde et al., 1972), three events have revolutionized its applicability *in vivo* which is of increased relevance to biologists: the availability and expansion of technologies for labeling biomolecules in solution or within a living cell (Bacia and Schwille, 2003; Miyawaki et al., 2003), the ability to perform fluctuation analysis within very small defined volumes and the efficient suppression of the background (Eigen and Rigler, 1994; Koppel et al., 1976; Rigler et al., 1993). Towards this purpose, the confocal setup was found by R. Rigler (Rigler and Widengren, 1990) to fulfill these requirements and to be ideal for investigating molecular processes within living cells *in situ* with single molecule sensitivity (Fig. 7), which has been shown to be possible for single organic dye molecules both in the solid state (Moerner and Kador, 1989; Orrit and Bernard, 1990) and in liquids (Bender et al., 1983b; Rigler and Mets, 1992; Shera et al., 1990). Ever since the discovery of Green Fluorescent Protein (GFP) from the jellyfish *Aequorea victoria* (Shimomura et al., 1962) and Red Fluorescent Protein (RFP) from various anthozoans (Fradkov et al., 2000; Matz et al., 1999; Wiedenmann et al., 2002), as well as the derivation of modified RFPs from *Discosoma sp.* (Shaner et al., 2004), the use of *in vivo* protein labeling has revolutionized life sciences and has made fluorescence microscopy a routine tool for the average biologist.

In order to perform temporal autocorrelation analysis the first step is to derive the intensity autocorrelation function:

$$C(\tau) = < I(t)I(t+\tau) >,$$

as an ensemble average of the measured intensity of the light *I(t)* at a certain time *t* and the intensity measured at a later time *I(t+τ)*. However, since the values of this equation depend on the properties of the experimental setup, the use of the normalized autocorrelation function is preferred:

$$G(\tau) = 1 + \frac{< \partial I(t)\partial I(t+\tau) >}{< I >^2}.$$

The use of the above formulation is more convenient since it is independent of the laser intensity, the efficiency of the detection method and the brightness of the used fluorophore (fluorescence quantum yield).

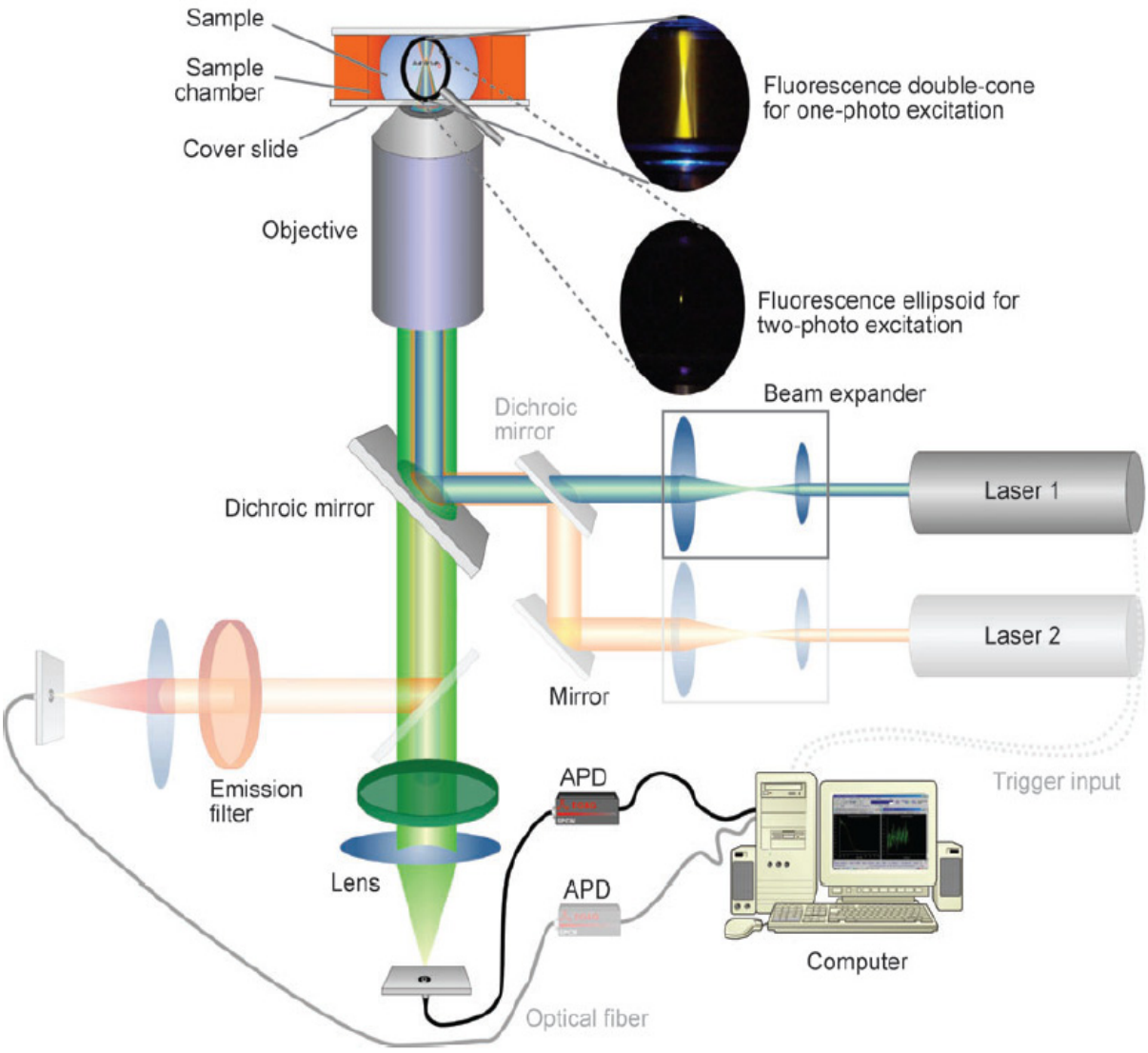

Fig. 7. **Schematic experimental setup for single-channel measurements (blue-green excitation and green detection channel) or dual-color cross-correlation (green and red beampaths).** The shape of the focal volume depends on the excitation mode (Haustein and Schwille, 2007).

Typically, in systems undergoing random fluorescence intensity fluctuations, $G(\tau)$ will vary around the mean value $G(\tau)=1$, but non-random processes will lead to the derivation of an autocorrelation curve, which contains information about the investigated system, such as the number of fluorescent molecules (and therefore the concentration of the molecule of interest in the sample), which is intrinsically determined from the variance of Poissonian processes (Ehrenberg and Rigler, 1974; Elson and Magde, 1974; Koppel, 1974). In a second step, fitting the values derived by the normalized autocorrelation curve to the most appropriate known autocorrelation functions that describes, for instance, free three-dimensional diffusion of the fluorophore without changes affecting its fluorescence:

$$G(\tau) = 1 + \frac{1}{N(1+\frac{\tau}{\tau_D})\sqrt{1+\frac{w_{xy}^2\tau}{w_z^2\tau_D}}},$$

or the free-diffusion of unbound molecules as compared to molecules bound to much larger particles:

$$G(\tau) = 1 + \frac{1}{N}\left(\frac{1-y}{\left(1+\frac{\tau}{\tau_{D_1}}\right)\sqrt{1+\frac{w_{xy}^2\tau}{w_z^2\tau_{D_1}}}} + \frac{y}{\left(1+\frac{\tau}{\tau_{D_2}}\right)\sqrt{1+\frac{w_{xy}^2\tau}{w_z^2\tau_{D_2}}}}\right),$$

allows the investigation of the presence or absence of molecular interactions in the sample. Details on these and other model autocorrelation functions, describing different processes, can be found for example in (Vukojevic et al., 2005).

Since FCS is used to determine translational diffusion of molecules, its advantage over other methods is that it is less sensitive to the size of molecules, since diffusion scales with the cubic root of the molecular weight – reviewed in (Zorrilla et al., 2008a). In addition, FCS can be readily performed on molecules in solution, which allows longer observation times, as compared to single-molecule detection on surfaces.

Among the first biologically relevant FCS measurements performed have been those of the study of the hybridization of nucleic acids (Kinjo and Rigler, 1995) and enzyme-substrate interactions (Rauer et al., 1996). However, numerous studies thereafter have expanded the applicability of FCS to DNA structure and interactions, protein function and interactions, protein structure, ligand-receptor binding reactions and even genetics – reviewed in (Vukojevic et al., 2005).

Focusing on transcription factor-DNA interactions and/or the derivation of their binding constants, relevant studies have been primarily performed in solution, using labeled recombinant proteins and oligonucleotides bearing the known binding sequences of transcription factors. Among the investigated transcription factors it is worth mentioning: the cro repressor of phage λ, the Lactose (Lac) repressor of *Escherichia coli* and the T7 RNA polymerase (Cook et al., 1990); the Tetracyclin (Tet) repressor and its binding to the Tet operator binding site (Chabbert et al., 1992); the prokaryotic transcriptional activator protein Nitrogen regulatory protein C (NtrC) (Rippe, 2000; Sevenich et al., 1998); p53 and its target DNA sequences

(Yakovleva et al., 2001); the binding of hexameric Replicative A (RepA) DNA-helicase to single stranded DNA (Xu et al., 2001); and the DNA-binding and intrinsic oligomerization of the yeast TATA binding protein (Tbp) (Khrapunov et al., 2002). Moreover, methods to carry out genome-wide gene expression analyses with single-molecule sensitivity (Korn et al., 2003) and the quantification of mRNA levels (Camacho et al., 2004) have been developed.

Finally, recent studies of transcription factor-DNA binding involve the investigation of Activator protein 1 (AP-1) and Nuclear Factor κB (NF-κB) (Kobayashi et al., 2004); Serum Response Factor (SRF) (Huet et al., 2005); the binding of Myocyte enhancer factor 2A (Mef2A) to its target DNA sequence (Octobre et al., 2005); plant transcription factors' Auxin Response Factors (ARFs) (Muto et al., 2006); the *Drosophila* Heat shock factor (Hsf) and its association with polytene DNA for the regulation of the *hsp70* genes after heat shock (Yao et al., 2006); the generic transcriptional activator Stimulatory protein 1 (Sp1) and its target GC-rich promoter sequences in essential and viral genes (Yeh et al., 2006); the Trans-acting transcription factor (Tat) of Human Immunodeficiency Virus (HIV) (Nandi et al., 2008); the control catabolite protein of gluconeogenic genes with its target operator sequences (Zorrilla et al., 2008b); and the interactions between the dimerization and binding domains of Vitellogenin binding protein (Vbp) and its target DNA sequences (Michelman-Ribeiro et al., 2009).

These studies – though performed in the majority of cases *in vitro* – underscore the importance of FCS as a method of choice for the study of molecular interactions in a non-destructive way. Especially in cases of systems involving interactions that have not been adequately described before (e.g. interactions of Hox transcription factors with their DNA binding sites *in vivo*) and thus cannot be reconstituted *in vitro*, the coupling of FCS to normal confocal microscopy allows to perform live sample analyses with higher penetration depth under physiological conditions (this study).

In the present work we use FCS with single-molecule sensitivity to study the interactions of Hox transcription factors with nuclear DNA in live cells and to describe their specific and non-specific binding to their potential target sites on the chromosome. Since the precise sequence and the number of binding sites of Hox transcription factors in the genome are largely unknown, we make use of active and inactive variants of the HD to extract information about specific versus non-specific

binding and derive by FCS their corresponding dissociation constants, a result hitherto not achieved by any other method of analysis.

RESULTS AND DISCUSSION

In this section, the functional analysis of synthetic Hox transcription factors (Antp and Scr) and the study of the interactions of synthetic Scr peptides with nuclear DNA in live cells, using quantitative imaging and FCS, are provided in the form of two published and one yet unpublished article. Moreover, our principal findings and the methodology used are discussed therein.

Function and specificity of synthetic Hox transcription factors in vivo

Dimitrios K. Papadopoulos[a], Vladana Vukojević[b], Yoshitsugu Adachi[c], Lars Terenius[b], Rudolf Rigler[d,e], and Walter J. Gehring[a,1]

[a]Department of Cell Biology, Biozentrum, University of Basel, 4056 Basel, Switzerland; [b]Department of Clinical Neuroscience, Karolinska Institutet, 17176 Stockholm, Sweden; [c]Department of Neuroscience, Institute of Psychiatry, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, United Kingdom; [d]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 17177 Stockholm, Sweden; and [e]Laboratory of Biomedical Optics, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland

Contributed by Walter J. Gehring, December 23, 2009 (sent for review November 3, 2009)

Homeotic (*Hox*) genes encode transcription factors that confer segmental identity along the anteroposterior axis of the embryo. However the molecular mechanisms underlying *Hox*-mediated transcription and the differential requirements for specificity in the regulation of the vast number of *Hox*-target genes remain ill-defined. Here we show that synthetic *Sex combs reduced* (*Scr*) genes that encode the Scr C terminus containing the homeodomain (HD) and YPWM motif (Scr-HD) are functional in vivo. Synthetic Scr-HD peptides can induce ectopic salivary glands in the embryo and homeotic transformations in the adult fly, act as transcriptional activators and repressors during development, and participate in protein-protein interactions. Their transformation capacity was found to be enhanced over their full-length counterpart and mutations known to transform the full-length protein into constitutively active or inactive variants behaved accordingly in the synthetic peptides. Our results show that synthetic *Scr*-HD genes are sufficient for homeotic function in *Drosophila* and suggest that the N terminus of Scr has a role in transcriptional potency, rather than specificity. We also demonstrate that synthetic peptides behave largely in a predictable way, by exhibiting *Scr*-specific phenotypes throughout development, which makes them an important tool for synthetic biology.

synthetic genes | transcriptional specificity | *Hox* genes | *Sex combs reduced* | homeotic transformations

Homeotic genes code for transcription factors that play an instrumental role in animal development by specifying the identity of body segments along the anteroposterior axis of the embryo (1–4). *Hox* genes have persisted in the animal kingdom; they are found in animals as diverse as worms and humans (5, 6) and the Homeodomain (HD), a helix-turn-helix DNA-binding domain, has been strikingly conserved in animals since before the bilaterian split (1, 7, 8). Sequence-specific binding of Hox proteins has been studied for the *Drosophila* Sex combs reduced (Scr) (9, 10), Antennapedia (Antp) (11, 12) and Ultrabithorax (Ubx) (11, 12) HDs. A consensus sequence TAATC/GC/G recognition core was identified in all of them, which alone is obviously not sufficient to confer transcriptional specificity, because it occurs statistically every kilobase in the genome. Similar sequence preferences have been identified for Deformed (Dfd) and Abdominal-B (Abd-B) (11), raising the question of how target specificity is achieved among different *Hox* paralogs.

A closer look into conserved residues outside the HD identified its amino-terminal YPWM motif that is present in almost all Hox proteins, from flies to vertebrates [with the exception of Abdominal-B (Abd-B), which has conserved only the tryptophan at position 3] (13). Extradenticle (Exd) and its mammalian homolog Pbx1 (14) were found to interact specifically with the YPWM motif of Hox proteins in vitro (15) and crystallographic analysis of a Ubx-Exd complex determined the topology of this interaction (16). Recently the link between the Antp YPWM motif and the transcriptional machinery was made through the identification of Bric-à-Brac interacting protein 2 (Bip2) (17). However, *Hox* cofactors do not suffice to entirely explain Hox specificity. The finding of cofactor independent Hox function (18) contributed to the realization that further sequences residing in the N terminus of Hox proteins might be the link for increased specificity in vivo.

In the present work we have derived synthetic *Scr* genes that encode the YPWM, HD and C terminus of the *Drosophila* Scr. The synthetic Scr-HD retained many functions of the full-length protein in vivo by participating in homeotic transformations, transcriptional regulation, and protein-protein interactions, thus reflecting to a great extent the properties of the native Scr protein. Constitutively active and inactive variants, in which threonine-6 and serine-7 of the Scr-HD have been substituted by alanines (Scr-HD_{AA}) or aspartates (Scr-HD_{DD}), respectively (19), behaved in the synthetic peptides similar as in the full-length protein, but the synthetic peptides exhibited stronger homeotic function as compared to their full-length counterpart. Moreover, we show that the synthetic peptides bind specifically to native Scr and consensus HD-binding sites in vitro and accumulate at sites of loose chromatin conformation in live salivary gland nuclei, where *Scr* is normally expressed during development (20). Taken together our results indicate that synthetic Scr peptides are functional in vivo and thus challenge the role of the N-terminal part of the endogenous fly peptide in transcriptional specificity.

Results

Synthetic *Scr* Genes are Capable of Inducing Ectopic Salivary Glands in the Embryo. *Scr* acts as a master control regulator of salivary gland morphogenesis during embryonic development (20–25). When expressed throughout the embryo, Scr is able to induce an additional pair of salivary glands anterior to parasegment 2 (20), where the normal salivary glands form. Posteriorly its function is restricted by *teashirt* (*tsh*) and *Abdominal-B* (*Abd-B*) (26). In accordance with these findings, the wild type (*Scr*-HD_{wt}) and the constitutively active (*Scr*-HD_{AA}) synthetic genes could also induce ectopic salivary glands in the embryonic head, while the constitutively inactive variant (*Scr*-HD_{DD}) could not (Fig. 1*A–C*) and displayed a normal pair of salivary glands, similar to the control embryos, also treated with heat shock (Fig. 1*D*).

Misexpression of Synthetic *Scr* Genes Causes Homeotic Transformations in the Adult Head. Both the fly *Scr* gene and its functional mouse homolog *Hox-a5* (previously described as *Hox-1.3*) have been shown to induce partial antenna-to-tarsus transformations in the fly head (27, 28). To examine the ability of Scr-HD_{wt} and Scr-

Author contributions: W.J.G. designed research; D.K.P., V.V., and Y.A. performed research; D.K.P. and V.V. contributed new reagents/analytic tools; D.K.P. analyzed data; and D.K.P., L.T., R.R., and W.J.G. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: walter.gehring@unibas.ch.

This article contains supporting information online at www.pnas.org/cgi/content/full/0914595107/DCSupplemental.

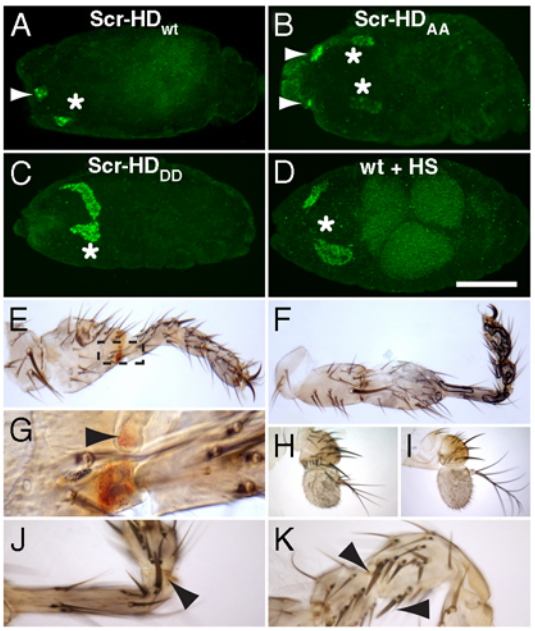

Fig. 1. Synthetic Scr peptides induce homeotic transformations in the fly (*A–C*). Expression of the synthetic genes throughout the embryo results in the formation of additional salivary glands in the cephalic region for *Scr*-HD_{wt} (*A*) and *Scr*-HD_{AA} (*B*), but not *Scr*-HD_{DD} (*C*). (*D*) Wild type embryo also treated with heat shock allows the development of one normal pair of salivary glands. Arrowheads show the ectopic—and asterisks, the normal—salivary glands. All constructs were induced using *Heat-Shock*-Gal4. Stainings are on stage 16 embryos for dCREB-A (19), a salivary gland luminal marker. Scale bar 100 μm. (*E–I*) Expression of *Scr*-HD_{wt} (*E*) and *Scr*-HD_{AA} (*F*) in the antennal disc results in complete antenna-to-tarsus transformations. (*G*) Magnification of the outlined area in (*E*) shows pigmented cells (*Arrowhead*) at the distal part of the transformed A3. (*H*) *Scr*-HD_{DD} only confers a small reduction in the size of the arista. (*I*) Wild type antenna. (*J–K*) Sex comb teeth on antennal tarsi generated by ectopic expression of *Scr*-HD_{wt} (*J*) and *Scr*-HD_{AA} (*K*) (*Arrowheads*). All transformations were generated using a *Dll*-Gal4 driver.

HD_{AA} to trigger homeotic transformations in vivo we expressed them ectopically in the antennal portion of the eye-antennal disc using the *Distalless* (*Dll*) enhancer (29). Transformation of the third antennal segment (A3) and the arista to a fully grown tarsus (Fig. 1*E* and *F*, respectively) was observed. The inactive mutant (Scr-HD_{DD}) could only induce a reduction in the size of the arista (Fig. 1*H*) as compared to the wt antenna (Fig. 1*I*). The ectopic tarsi generated by gain-of-function of *Scr*-HD_{wt} displayed a patch of red pigmentation, an indication that a group of cells might have been transformed to ectopic eye-cells (Fig. 1*G*). Moreover, the presence of one to two sex comb teeth on the ectopic tarsi mediated by Scr-HD_{wt} (Fig. 1*J*) and Scr-HD_{AA} (Fig. 1*K*) indicated that the ectopic tarsi are prothoracic (T1), normally specified by *Scr* (28).

Synthetic Scr Peptides Participate in Protein-Protein Interactions with Pax Transcription Factors in Vivo. At the molecular level, Hox transcription factors are known to participate in protein-protein interactions. Such interactions have been demonstrated to take place between two HDs or a HD and a PAIRED domain (PD), resulting in mutual inhibition of DNA-binding properties, which leads to defects that resemble mutant phenotypes, caused by absence of either gene product. A phenotypic manifestation of such a negative posttranslational regulatory inhibition is the eye reduction caused by ectopic expression of various Hox proteins in the eye-antennal disc, shown to be triggered by binding of the HD of Hox proteins to the PD of eyeless (ey) (30, 31). In accordance with these findings ectopic expression of *Scr*-HD_{wt} and *Scr*-HD_{AA} using dpp^{blink}-Gal4 exhibited strong eye phenotypes (Fig. 2*A–D*) as compared to the full-length peptide (30), ranging from reduction to the complete absence of eyes. No abnormal eye phenotype

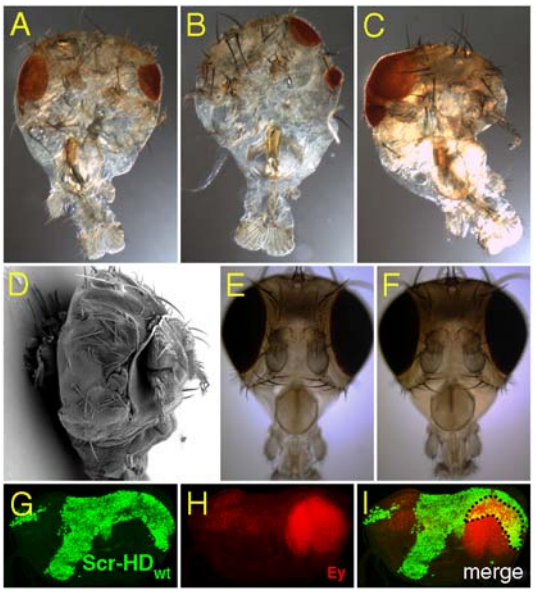

Fig. 2. Synthetic Scr peptides interact with Pax transcription factors in vivo (*A–D*). Eye-reduction phenotypes exhibited by ectopic expression of *Scr*-HD_{wt} (*A*, *B*, and *D*) and *Scr*-HD_{AA} (*C*). According to the strength of expression, different lines exhibited phenotypes ranging from eye-reduction (*A–C*) to eye-absence (*D*). (*E*) Ectopic expression of *Scr*-HD_{DD} resulted in no detectable phenotype. (*F*) Wild type head. (*G–I*) Ectopic expression of *Scr*-HD_{wt} in the eye-disc (*G*) does not repress *ey* (*H*). Dashed lines show the domain of colocalization of Scr-HD_{wt} and Ey (*I*). The dpp^{blink}-Gal4 driver has been used throughout.

was observed with *Scr*-HD_{DD} (Fig. 2*E*), and the eyes in this case were indistinguishable from the control (Fig. 2*F*). Colocalization of Scr-HD_{wt} (Fig. 2*G*) with endogenous Ey (Fig. 2*H*) suggested their interaction at the posttranslational level in the region of coexpression (Fig. 2*I*, *Dashed Line*).

Synthetic Scr-HD Peptides Act as Transcriptional Activators and Repressors in the Eye-Antennal Disc. If Scr-HD peptides can cause antenna-to-tarsus transformations, genes responsible for antennal development must be repressed and leg determination genes must be activated in the antennal disc (32). Ectopic expression of *Antp* or *Scr* in the antennal disc is able to repress *Spalt major* (*Salm*) at the transcriptional level, thus preventing the differentiation to an antenna and allowing the initiation of the leg determination program (21, 33). In agreement with this principle, both Scr-HD_{wt} (Fig. 3*A* and *D*) and Scr-HD_{AA} (Fig. 3*B* and *E*) repressed *Salm*. As expected, no repression was observed with Scr-HD_{DD} (Fig. 3*C* and *F*).

The same paradigm was found to apply for *distal antenna* (*dan*), another antenna determination gene normally expressed in the eye-antennal but not in the leg disc (34). Ectopic *Antp* in the antennal portion of the disc, induced in a mutant *spineless* (*ss*) background, results in the repression of *dan*, thus transforming the antenna into leg structures, a result suggesting that *dan* is repressed in discs undergoing tarsal transformations (34). Our findings show that this also occurs with *Scr*. Scr-HD_{wt} partially repressed *dan* (Fig. 3*G*, arrowhead, and *J*). In the case of Scr-HD_{AA} repression of *dan* was complete (Fig. 3*H* and *K*), while no repression was detected with Scr-HD_{DD} (Fig. 3*I* and *L*).

Similar behavior was expected to apply in the repression of *Homothorax* (*Hth*), the function of which in the antennal determination program has been described extensively (35). Although expressed both in the antennal and the leg disc during development, its expression domain in the antennal disc largely overlaps with *Dll*, whereas in the leg disc the two gene products are distributed to distinct, nonoverlapping regions (36, 37).

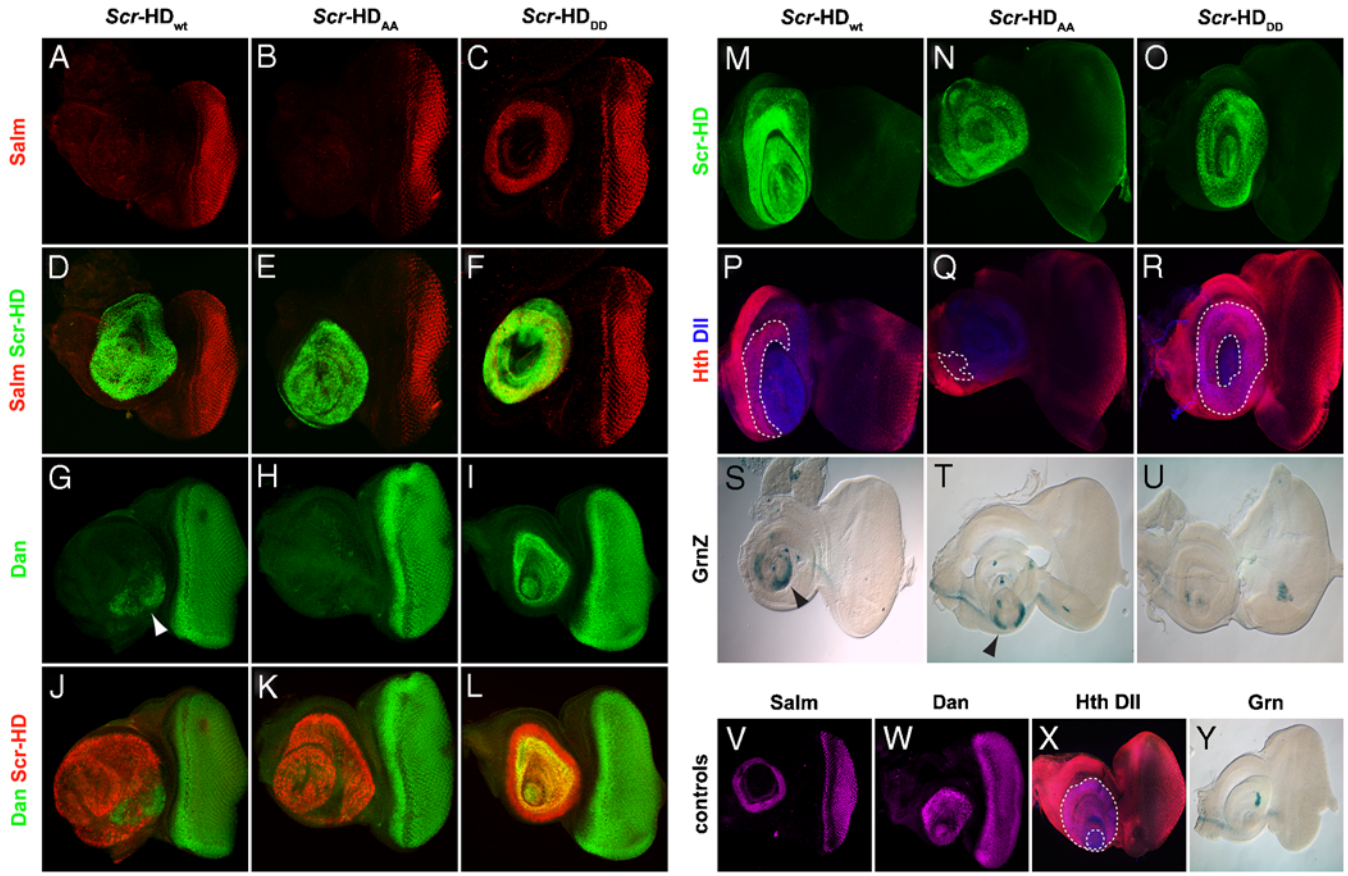

Fig. 3. Synthetic Scr-HD peptides act as transcriptional activators and repressors in the antennal primordium. (*A–F*) Repression of *Spalt major* (*Salm*) in the antennal disc by ectopic expression of *Scr*-HD$_{wt}$ (*A* and *D*) and *Scr*-HD$_{AA}$ (*B* and *E*). No repression was observed with *Scr*-HD$_{DD}$ (*C* and *F*). (*G–L*) Repression of *distal antenna* (*dan*) in the antennal disc is complete with Scr-HD$_{AA}$ (*H* and *K*) and incomplete with Scr-HD$_{wt}$ (*G* and *J*), leaving a patch of cells that retain Dan activity (*Arrowhead* in *G*). These cells do not express ectopic *Scr*-HD$_{wt}$. (*I* and *L*) Scr-HD$_{DD}$ does not repress *dan*. (*M–R*) Repression of *Hth* results in a shift of the Hth-Dll boundary. Partial repression of *Hth* by Scr-HD$_{wt}$ (*M* and *P*) and Scr-HD$_{AA}$ (*N* and *Q*), as compared to Scr-HD$_{DD}$ (*O* and *R*), where no repression is observed. (*S–U*) X-gal stainings showing the activation of *grain* (*grn*) by Scr-HD$_{wt}$ (*S*) and Scr-HD$_{AA}$ (*T*) (*Arrowheads*). Scr-HD$_{DD}$ (*U*) fails to activate ectopic *grn* expression. (*V–Y*) Normal expression of *Salm* (*V*), *dan* (*W*), *Hth-Dll* (*X*), and *grn* (*Y*) in eye-antennal discs. *Dll*-Gal4 was used to drive expression of all constructs (*A-Y*).

Because ectopic Antp in the antennal disc represses *Hth* (36) we assumed a general mechanism, according to which tarsal transformations in the antenna repress and thus restrict *Hth* outside of the *Dll* expression domain, resulting in a shift of the Hth-Dll boundary. Indeed, ectopic expression of *Scr*-HD$_{wt}$ (Fig. 3*M* and *P*) and *Scr*-HD$_{AA}$ (Fig. 3*N* and *Q*) resulted in partial or complete repression of *Hth* respectively, while no repression was seen with the inactive construct *Scr*-HD$_{DD}$ (Fig. 3*O* and *R*).

Inversely, to probe the transformed antennal discs for ectopically activated leg-specific genes, we tested the synthetic Scr-HD variants in their ability to activate *grain* (*grn*). *grn* encodes a GATA transcription factor (GATAc) that plays an important role in cell rearrangement during morphogenesis. It is expressed in the Central Nervous System (CNS), midgut, and lateral ectoderm during development (38). *grn* is normally expressed in the leg disc but not in the antennal disc; however, it is activated in the antennal disc upon ectopic expression of *Antp* [*grn* enhancer trap lines have been published as rK781 in (33) and as klecks in (39)]. Fig. 3*S–U* show the ectopic activation of *grn* mediated by the wild type and active constructs, whereas Scr-HD$_{DD}$ failed to activate transcription of the reporter (Fig. 3*U*). Taken together, these results demonstrate that the synthetic Scr peptides participate, directly or indirectly, in both transcriptional repression and activation in vivo.

Cells that Maintain *dan* Activity in the Antennal Primordium Give Rise to Ectopic Eyes on the Antennal Tarsi. We were next interested to provide an explanation as to whether the pigmented cells observed in the transformed antennae (Fig. 1*B* and *C*) are indeed eye structures. We hypothesized that cells that retained *dan* activity in the antennal disc (Fig. 3*G*, *Arrowhead*) might account for the formation of ectopic eyes on the transformed tarsi upon misexpression of *Scr*-HD$_{wt}$. Therefore, we tested the presence of photoreceptor cell markers in Dan protein positive cells in the antennal disc (Fig. 4*A–D*). Ectopic expression of *Embryonic lethal abnormal visual system* (*Elav*) colocalized with cells that still expressed *dan* but not ectopic *Scr*-HD$_{wt}$ (Fig. 4*D*), suggesting that these cells have acquired neuronal fate, similar to the differentiated photoreceptors in the eye portion of the disc. *dan* participates in both antennal and eye development (34, 40), and because these cells cannot differentiate into an antenna, they are able to form ectopic eyes in the transformed A3 segment. Scanning Electron Microscopy (SEM) of adult antennae confirmed that the pigmented cells observed previously (Fig. 1*B* and *C*) are indeed compound eyes, comprising several ommatidia (Fig. 4*E*).

Molecular Interactions of Synthetic Scr Peptides in Live Salivary Gland Cells Using High-Resolution Quantitative APD Imaging. In order to understand at the molecular level differences between the active and inactive peptide variants in their ability to interact with nuclear DNA, we studied their interactions with chromosomal DNA in live cells using advanced fluorescence imaging with Avalanche Photo Detectors (APD imaging) (41). Here we substantiate the phenotypic findings by molecular imaging. Elaborate quantitative study of Scr-HD–DNA interactions in live salivary gland cells is presented in the following paper (42).

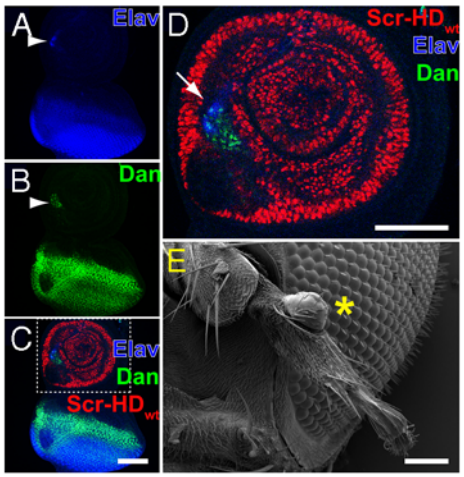

Fig. 4. Cells with no ectopic Scr-HD$_{wt}$ activity fail to repress *dan* but activate *Elav* and thus differentiate into compound eyes. (*A*) *Elav* gain-of-function, (*B*) Dan (*Arrowheads*), and (*C*) merge of (*A*) and (*B*) with Scr-HD$_{wt}$. (*D*) Magnification of outlined area in (*C*). Arrow points at cells that express both *dan* and *Elav* but not ectopic *Scr*-HD$_{wt}$. (*E*) Adult antenna transformed into a tarsus, which bears a small ectopic eye in the A3. Note the presence of ommatidia and interommatidial bristles (*Red Asterisk*). Scale bars in (*C*) 100 μm and in (*D*, *E*) 50 μm.

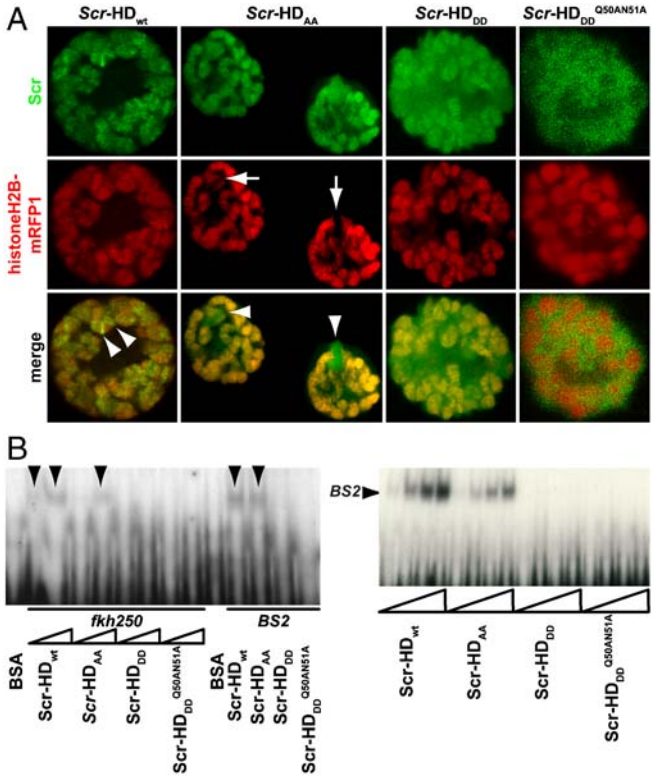

Fig. 5. High-resolution APD imaging of DNA-Scr-HD interactions in live cells. (*A*) Third instar salivary gland polytene nuclei expressing *Scr*-HD$_{wt}$, *Scr*-HD$_{AA}$, *Scr*-HD$_{DD}$, and *Scr*-HD$_{DD}^{Q50AN51A}$ under the control of *dpp*^{blink}-Gal4. Ubiquitously expressed mRFP1-tagged histone H2B was used to visualize chromatin. Scr-HD$_{wt}$ and Scr-HD$_{AA}$ readily associate with the chromosomes (as shown in the *Green Channel*) but also show sites of accumulation along the chromosome where loose chromatin compaction is shown as a low histone signal (*Arrows*). Arrowheads point at sites of high accumulation observed for Scr-HD$_{wt}$ and Scr-HD$_{AA}$. The nucleus expressing the inactive *Scr*-HD$_{DD}$ shows some association of the transcription factor with the DNA, but it is also dispersed in the nucleoplasm. There is no pronounced banding pattern observed in this case, which suggests absence of specific binding. Scr-HD$_{DD}^{Q50AN51A}$ appears almost completely excluded from the chromosomes, mainly residing in the nucleoplasm. Scale bars in all cases are 20 μm. (*B*) Electrophoretic Mobility Shift Assay (EMSA) shows that only Scr-HD$_{wt}$ and Scr-HD$_{AA}$ bind DNA specifically in vitro. Both variants bound more strongly to *BS2* than *fkh250* (*Left*). Titration of peptide concentration (*Right*) revealed that even at high concentrations of transcription factor, Scr-HD$_{DD}$ and Scr-HD$_{DD}^{Q50AN51A}$ do not bind DNA (*BS2*) specifically.

To visualize chromatin, a ubiquitously expressed *histone-H2B-mRFP1* line was used. *mCitrine-Scr-HD* fusions expressed in salivary gland polytene nuclei showed that Scr-HD$_{wt}$ and Scr-HD$_{AA}$ peptides associate significantly with the DNA (Fig. 5). As observed for Scr-HD$_{wt}$ and Scr-HD$_{AA}$, the transcription factor did not associate uniformly with the DNA but accumulated at sites of loose chromatin conformation where the histone signal was almost absent (Fig. 5*A*, first and second column). These sites should correspond to transcriptionally active regions. In contrast, Scr-HD$_{DD}$ showed some association with chromatin, but the transcription factor was also observed in the nucleoplasm among polytene chromosomes (Fig. 5*A*, third column). To understand this weaker yet substantial association of Scr-HD$_{DD}$ with chromatin we additionally mutated residues 50 and 51 of the Scr-HD to alanines (*Scr*-HD$_{DD}^{Q50AN51A}$). Glutamine-50 and asparagine-51 participate in third helix binding of the Antp HD to the DNA (43). We assumed that these mutations would behave accordingly in the synthetic Scr-HD$_{DD}$ peptide, acting in synergy with T6D and S7D and rendering the peptide interactions with chromatin even weaker. As expected, the synthetic Scr-HD$_{DD}^{Q50AN51A}$ peptides appeared markedly dispersed in the nucleoplasm (Fig. 5*A*, fourth column) and excluded from chromatin. Flies expressing *Scr*-HD$_{DD}^{Q50AN51A}$ constitutively in the embryo, or ectopically in the antennal disc, did not display any of the embryonic or adult phenotypes.

Synthetic Scr-HD Peptides Bind to DNA Specifically in Vitro. To further investigate the observed association of Scr-HD$_{DD}$ with chromatin, we studied the specific binding of the synthetic peptides to DNA in vitro by gel-shift assays (Fig. 5*B*). Using the native Scr binding site *fkh250* and the generic HD-binding site *BS2*, we observed specific binding for both Scr-HD$_{wt}$ and Scr-HD$_{AA}$ but not for the inactive peptides Scr-HD$_{DD}$ and Scr-HD$_{DD}^{Q50AN51A}$, not even at high peptide concentrations (Fig. 5*B*, *Left*). Binding to *BS2* appeared to be stronger for the transcriptionally active variants (Fig. 5*B*, *Right*) and was used for the titration. This result suggested that binding of Scr-HD$_{DD}$ to chromosomal DNA in polytene nuclei is largely nonspecific, because Scr-HD$_{DD}$ has additionally no homeotic function in the fly. Thus, the imaging analysis in live cells and the in vitro binding assay support our genetic findings and strengthen the notion that synthetic *Scr* genes function in a predictable way in vivo.

Discussion

The genetic role of *Hox* genes has been studied extensively by gain- and loss-of-function experiments (24, 27, 35, 44–46), and the properties of the HD-DNA complex have been elucidated in solution by NMR (43). Nevertheless, the precise mechanisms that orchestrate the transcriptional regulation of the vast number of *Hox* target genes remain elusive. Here we have derived synthetic *Scr* genes and analyzed their function throughout development. We demonstrated that they are able to induce homeotic transformations in the adult fly and embryo, repress and activate antenna and leg-specific genes, respectively, and participate in protein-protein interactions. Our results support that they are functional in vivo and thus question the role of the N terminus of Scr in transcriptional specificity.

Hox-mediated antenna-to-tarsus transformations proceed through repression of genes necessary for antennal specification and ectopic activation of leg-specific genes in the antennal disc.

This transformation is DNA-binding-dependent for Antp (30), because residues that impair binding of the HD to the minor groove (47) also abolish its transformation capacity in vivo. This suggested a similar requirement for Scr and thus for synthetic Scr peptides. Indeed, the latter bound putative Scr and generic HD-binding sites in vitro (Fig. 4*B*), repressed antennal-specific genes (*Salm*, *dan*, and *Hth*) and activated leg-specific genes (*grn*) in the antennal disc (Fig. 3). The inability of Scr-HD_{DD} to trigger any of these phenotypes is in line with substitutions of threonine 6 and serine 7 to aspartates in the full-length protein, which impaired its DNA-binding activity in vitro (19). Our findings suggest that these mutations also abolish the capacity of Scr-HD_{DD} to participate in HD-PD or HD-HD interactions. Although mutation of glutamate 19 to glycine abolished the dimerization capacity of Antp in vivo (30), in the case of Scr other residues of the HD, in addition to glutamate 19, might participate in protein-protein interactions, or, alternatively, the negative charge introduced by the aspartates might be responsible for abolishing these interactions through electrostatic repulsion. The fact that Scr-HD_{AA} exhibited similar phenotypes as Scr-HD_{wt} (compare Fig. 2*C* to Fig. 2*A*, *B* and *D*) is in favor of this scenario.

Surprisingly, the homeotic transformations observed in the adult fly (Fig. 1*E–I*) were considerably stronger than the ones observed with the full-length protein (Fig. 6*A* and *C*). Partial antenna-to-tarsus transformations were obtained with both the dpp^{blink} (Fig. 6*A*, as compared to Fig. 4*E*) and the *Dll* enhancers (Fig. 6*C* and *D*, as compared to Fig. 1*E*) and the eye-reduction phenotypes were comparable in strength to the full-length protein (compare Fig. 6*B* to Fig. 2*A*, *B*, and *D*). The same applied in the repression of antennal-specific markers (*Salm* in Fig. 6*E* and *F*, *dan* in Fig. 6*G* and *H*, *Dll* in Fig. 6*I* and *J*, and *Hth* in Fig. 6*K* and *L*) and the greater overlap observed between Hth and Dll (Fig. 6*M*, areas indicated by white dashed lines, as compared to Fig. 3*P*) supported the partiality of the tarsal transformation. Finally, ectopic expression of *Elav* on the antennal disc was found to colocalize with Dan protein positive cells (Fig. 6*N–P*). Weaker transformations obtained by the full-length Scr (27, 28) support the notion that the synthetic genes exhibit stronger homeotic function in vivo. In the paradigm of salivary gland induction, a morphogenetic process initiated by *Scr*, the synthetic genes behaved in a predictable way. They triggered the formation of an additional pair of salivary glands in the region of the embryonic head (Fig. 1*A*), comparable to those induced by the full-length protein (19).

Four lines of evidence support that the function of the synthetic genes is specific rather than a generic HD effect. First, the sex comb teeth observed on the antennal tarsi indicate prothoracic leg identity (T1), specified by *Scr*. Second, the ectopic salivary glands in the embryo imply *Scr*-specific function (26). Third, mutations in the HD of the full-length peptide (19) behaved accordingly in synthetic peptides. Finally, specificity of Antp and Scr is owed to the N-terminal part of the HD (44), also present in the synthetic peptides.

It has been proposed that the large N-terminal part of Ubx and Scr is required for transcriptional activation in vivo (48). Our data show that this need not be the case for several *Scr*-specific functions. So, if *Scr*-specificity is achieved by synthetic genes, what is the function of the large N-terminal sequence of Scr in vivo? Sequence comparison of the fly Scr with its insect and vertebrate homologs (Fig. S1) revealed the divergence of the N terminus of the protein in length and sequence, with the exception of the MSSYFVNS, the DYTQL and the SCKYA motifs. Both the fly (48) and the murine MSSYFVNS domain in flies (49) seem to be limited to contributing to transcriptional potency rather than specificity, because its deletion resulted in homeotic transformations (49) and ectopic activation of *Scr* target genes, albeit weaker than the wild type protein (48, 49). Substitution of serine-10 by leucine in the Scr^{14} hypomorphic allele only resulted in a mild decrease in the number of sex comb teeth in the adult (50). The same applies to the DYTQL motif, because Scr^{15}, another hypomorphic allele of *Scr* lacking this motif, displayed partially compromised *Scr*-activity but no loss of specificity in vivo (50). When parts of the N terminus of the murine Hox-a5 were removed, its transcriptional activity in live cells and its DNA-binding in vitro were considerably weaker, though not completely abolished (49). Although Hox-a5 lacking the complete N terminus could not trigger homeotic transformations in the fly (49), this indicates that the N terminus is responsible for regulating transcriptional levels. Our results show that the HD, YPWM motif and C terminus of Scr are both necessary and sufficient for providing transcriptional specificity in vivo. A change in potency, rather than in specificity, was also observed for Antp lacking fractions of its N terminus (44), suggesting a general requirement of the latter for transcriptional "fine-tuning" rather than specificity among *Hox* paralogs and homologs. An example of a dose-dependent developmental output of *Hox* gene products has already been described for Ubx-mediated repression of *Dll* (48) and might apply to a few Hox peptides. Quantitative studies might help to precisely describe such features of Hox peptides in vivo.

Our work also supports the hypothesis that synthetic genes and/or peptides behave to a large extent in a predictable way. This notion is central to synthetic biology (51–53), not only for engineering artificial processes, but also for developing new techniques for gene and peptide therapy. In the long run, functional synthetic genes/peptides may bear advantages for medical

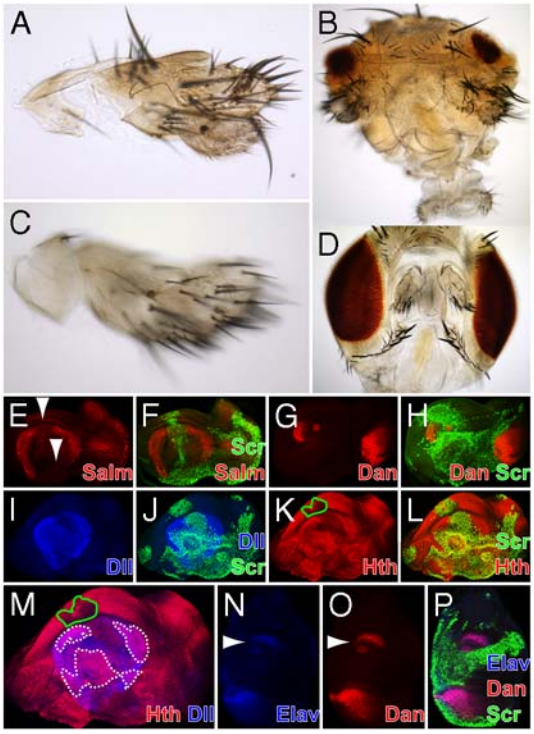

Fig. 6. The full-length Scr peptide exhibits weaker homeotic function than its synthetic counterparts in vivo. (*A*) Partial antenna-to-tarsus transformation mediated by Scr full-length using the dpp^{blink} enhancer. (*B*) The same gain-of-function results in eye-reduction in the head. (*C*) Transformation using the *Dll* enhancer is also incomplete. (*D*) Adult head of a fly expressing Scr full-length using *Dll*-Gal4. (*E–F*) Repression of *Salm* by the full-length Scr in the antennal disc (*Arrowheads* in *E*) is complete. (*G–H*) The same applies in the repression of *dan*. (*I–M*) Repression of *Dll* and *Hth* is incomplete (*I*, *K*, and *M*) and considerable overlap of Hth and Dll is observed in the antennal disc (*Dashed White Lines* in *M*). Scr represses *Hth* also outside the *Dll* expression domain (*Solid Green Line* in *K* and *M*). (*N–P*) Cells that maintain Dan activity (*Arrowhead* in *O*) in the antennal disc express ectopic *Elav* (*Arrowhead* in *N*). The full-length *Scr* has been induced by dpp^{blink}-Gal4 (*E–P*).

applications. Due to their size, they are expected to be considerably easier to engineer/synthesize and have better penetration efficiency into target cells or tissue—one of the greatest obstacles in peptide therapy. More studies in this direction will help unravel such perspectives. In the following paper (42), we take first steps in studying quantitatively the interactions of synthetic Scr peptides with DNA in live cells.

Materials and Methods

Cloning Procedure and Fly Transgenesis. Generation of plasmids was performed using standard procedures. Transgenic lines were generated as described (54). Fly stocks used are outlined in *SI Materials and Methods*.

Immunohistochemistry. Antibody stainings were performed as outlined in *SI Materials and Methods*.

High-Resolution APD Imaging. High-resolution APD imaging was performed on a uniquely modified ConfoCor3 instrument (Carl Zeiss, Jena, Germany) (refer to *SI Materials and Methods* for more detail).

Electrophoretic Mobility Shift Assay (EMSA). All *mCitrine-Scr-HD* fragments were cloned to pCR2.1-TOPO Vector (Invitrogen) and the peptides produced in vitro using TNT® T7 Quick Coupled Transcription/Translation System (Promega). EMSA was then performed as described (31).

ACKNOWLEDGMENTS. This work was supported by the Kantons of Basel-Stadt and Basel-Landschaft, a grant from the Swiss National Science Foundation, the European Network of Excellence "Cells into Organs," the Swedish Brain Foundation, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, and the Ministry of Sciences and Technological Development of Serbia (Grants no. 142025 and 142019).

1. Gehring WJ (1987) Homeo boxes in the study of development. *Science* 236(4806):1245–1252.
2. Lewis EB (1978) A gene complex controlling segmentation in Drosophila. *Nature* 276(5688):565–570.
3. Mann RS, Morata G (2000) The developmental and molecular biology of genes that subdivide the body of Drosophila. *Annu Rev Cell Dev Biol* 16:243–271.
4. McGinnis W, Krumlauf R (1992) Homeobox genes and axial patterning. *Cell* 68(2):283–302.
5. Kappen C, Ruddle FH (1993) Evolution of a regulatory gene family: HOM/HOX genes. *Curr Opin Genet Dev* 3(6):931–938.
6. Kmita-Cunisse M, Loosli F, Bierne J, Gehring WJ (1998) Homeobox genes in the ribbonworm Lineus sanguineus: Evolutionary implications. *Proc Natl Acad Sci USA* 95(6):3030–3035.
7. Affolter M, Schier A, Gehring WJ (1990) Homeodomain proteins and the regulation of gene expression. *Curr Opin Cell Biol* 2(3):485–495.
8. Scott MP, Tamkun JW, Hartzell GW 3rd (1989) The structure and function of the homeodomain. *Biochim Biophys Acta* 989(1):25–48.
9. Ryoo HD, Mann RS (1999) The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* 13(13):1704–1716.
10. Odenwald WF, Garbern J, Arnheiter H, Tournier-Lasserve E, Lazzarini RA (1989) The Hox-1.3 homeo box protein is a sequence-specific DNA-binding phosphoprotein. *Genes Dev* 3(2):158–172.
11. Ekker SC, et al. (1994) The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. *EMBO J* 13(15):3551–3560.
12. Muller M, et al. (1988) Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J* 7(13):4299–4304.
13. Qian YQ, et al. (1992) NMR structure determination reveals that the homeodomain is connected through a flexible linker to the main body in the Drosophila Antennapedia protein. *Proc Natl Acad Sci USA* 89(22):10738–10742.
14. Phelan ML, Rambaldi I, Featherstone MS (1995) Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol Cell Biol* 15(8):3989–3997.
15. Johnson FB, Parker E, Krasnow MA (1995) Extradenticle protein is a selective cofactor for the Drosophila homeotics: Role of the homeodomain and YPWM amino acid motif in the interaction. *Proc Natl Acad Sci USA* 92(3):739–743.
16. Passner JM, Ryoo HD, Shen L, Mann RS, Aggarwal AK (1999) Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* 397(6721):714–719.
17. Prince F, et al. (2008) The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* 135(9):1669–1679.
18. Galant R, Walsh CM, Carroll SB (2002) Hox repression of a target gene: Extradenticle-independent, additive action through multiple monomer binding sites. *Development* 129(13):3115–3126.
19. Berry M, Gehring W (2000) Phosphorylation status of the SCR homeodomain determines its functional activity: Essential role for protein phosphatase 2A,B'. *EMBO J* 19(12):2946–2957.
20. Panzer S, Weigel D, Beckendorf SK (1992) Organogenesis in Drosophila melanogaster: Embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. *Development* 114(1):49–57.
21. Abzhanov A, Holtzman S, Kaufman TC (2001) The Drosophila proboscis is specified by two Hox genes, proboscipedia and Sex combs reduced, via repression of leg and antennal appendage genes. *Development* 128(14):2803–2814.
22. Andrew DJ, Henderson KD, Seshaiah P (2000) Salivary gland development in Drosophila melanogaster. *Mech Dev* 92(1):5–17.
23. Haberman AS, Isaac DD, Andrew DJ (2003) Specification of cell fates within the salivary gland primordium. *Dev Biol* 258(2):443–453.
24. Percival-Smith A, Weber J, Gilfoyle E, Wilson P (1997) Genetic characterization of the role of the two HOX proteins, Proboscipedia and Sex Combs Reduced, in determination of adult antennal, tarsal, maxillary palp and proboscis identities in Drosophila melanogaster. *Development* 124(24):5049–5062.
25. Kerman BE, Cheshire AM, Andrew DJ (2006) From fate to function: The Drosophila trachea and salivary gland as models for tubulogenesis. *Differentiation* 74(7):326–348.
26. Andrew DJ, Horner MA, Petitt MG, Smolik SM, Scott MP (1994) Setting limits on homeotic gene function: Restraint of Sex combs reduced activity by teashirt and other homeotic genes. *EMBO J* 13(5):1132–1144.
27. Yao LC, Liaw GJ, Pai CY, Sun YH (1999) A common mechanism for antenna-to-Leg transformation in Drosophila: Suppression of homothorax transcription by four HOM-C genes. *Dev Biol* 211(2):268–276.
28. Zhao JJ, Lazzarini RA, Pick L (1993) The mouse Hox-1.3 gene is functionally equivalent to the Drosophila Sex combs reduced gene. *Genes Dev* 7(3):343–354.
29. Calleja M, Moreno E, Pelaz S, Morata G (1996) Visualization of gene expression in living adult Drosophila. *Science* 274(5285):252–255.
30. Plaza S, et al. (2008) Cross-regulatory protein-protein interactions between Hox and Pax transcription factors. *Proc Natl Acad Sci USA* 105(36):13439–13444.
31. Plaza S, et al. (2001) Molecular basis for the inhibition of Drosophila eye development by Antennapedia. *EMBO J* 20(4):802–811.
32. Gehring WJ, Kloter U, Suga H (2009) Evolution of the Hox gene complex from an evolutionary ground state. *Curr Top Dev Biol* 88:35–61.
33. Wagner-Bernholz JT, Wilson C, Gibson G, Schuh R, Gehring WJ (1991) Identification of target genes of the homeotic gene Antennapedia by enhancer detection. *Genes Dev* 5(12b):2467–2480.
34. Emerald BS, Curtiss J, Mlodzik M, Cohen SM (2003) Distal antenna and distal antenna related encode nuclear proteins containing pipsqueak motifs involved in antenna development in Drosophila. *Development* 130(6):1171–1180.
35. Casares F, Mann RS (2001) The ground state of the ventral appendage in Drosophila. *Science* 293(5534):1477–1480.
36. Casares F, Mann RS (1998) Control of antennal versus leg development in Drosophila. *Nature* 392(6677):723–726.
37. Dong PD, Chu J, Panganiban G (2000) Coexpression of the homeobox genes Distal-less and homothorax determines Drosophila antennal identity. *Development* 127(2):209–216.
38. Brown S, Castelli-Gair Hombria J (2000) Drosophila grain encodes a GATA transcription factor required for cell rearrangement during morphogenesis. *Development* 127(22):4867–4876.
39. Grieder NC, Morata G, Affolter M, Gehring WJ (2009) Spalt major controls the development of the notum and of wing hinge primordia of the Drosophila melanogaster wing imaginal disc. *Dev Biol* 329(2):315–326.
40. Curtiss J, Burnett M, Mlodzik M (2007) Distal antenna and distal antenna-related function in the retinal determination network during eye development in Drosophila. *Dev Biol* 306(2):685–702.
41. Vukojevic V, et al. (2008) Quantitative single-molecule imaging by confocal laser scanning microscopy. *Proc Natl Acad Sci USA* 105(47):18176–18181.
42. Vukojević V, Papadopoulos DK, Terenius L, Gehring WJ, Rigler R Quantitative study of synthetic Hox transcription factor-DNA interactions in live cells. *Proc Natl Acad Sci USA* doi:10.XXXX/XXXXXXXXXX.
43. Gehring WJ, et al. (1994) Homeodomain-DNA recognition. *Cell* 78(2):211–223.
44. Gibson G, Schier A, LeMotte P, Gehring WJ (1990) The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* 62(6):1087–1103.
45. Morata G, Sanchez-Herrero E (1999) Patterning mechanisms in the body trunk and the appendages of Drosophila. *Development* 126(13):2823–2828.
46. Wakimoto BT, Kaufman TC (1981) Analysis of larval segmentation in lethal genotypes associated with the antennapedia gene complex in Drosophila melanogaster. *Dev Biol* 81(1):51–64.
47. Otting G, et al. (1990) Protein–DNA contacts in the structure of a homeodomain–DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J* 9(10):3085–3092.
48. Tour E, Hittinger CT, McGinnis W (2005) Evolutionarily conserved domains required for activation and repression functions of the Drosophila Hox protein Ultrabithorax. *Development* 132(23):5271–5281.
49. Zhao JJ, Lazzarini RA, Pick L (1996) Functional dissection of the mouse Hox-a5 gene. *EMBO J* 15(6):1313–1322.
50. Sivanantharajah L, Percival-Smith A (2009) Analysis of the sequence and phenotype of Drosophila Sex combs reduced alleles reveals potential functions of conserved protein motifs of the Sex combs reduced protein. *Genetics* 182(1):191–203.
51. Benner SA, Sismour AM (2005) Synthetic biology. *Nat Rev Genet* 6(7):533–543.
52. Kaznessis YN (2007) Models for synthetic biology. *BMC Syst Biol* 1(47).
53. Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U (2006) Plasticity of the cis-regulatory input function of a gene. *PLoS Biol* 4(4):e45.
54. Spradling AC, Rubin GM (1982) Transposition of cloned *P* elements into Drosophila germ line chromosomes. *Science* 218(4570):341–347.

Supporting Information

Papadopoulos et al. 10.1073/pnas.0914595107

SI Text

Materials and Methods. ***Cloning Procedure and Fly Transgenesis.*** The fly stocks used were: *CyO/Sp*; *dpp*^{blink}*-Gal4/TM6B* (1); *ubi-histoneH2B-mRFP1/CyO*; *Dr/TM6B* (a kind gift of Yohanns Bellaïche); *Dll-Gal4(MD23)/CyO*; *Dr/TM6B* and *Dll-Gal4(MD713)/Dll-Gal4(MD713)* (2); *CyO/Sp*; *heat-shock-Gal4* (Bloomington stock center, stock number 1799); *CyO/Sp*; *grn-lacZ/TM6B* (3, 4); *UAS-Scr*; *fr10/TM3* (a kind gift of Deborah Andrew); and *Dll-Gal4(MD23)/CyO*; *grn-lacZ/TM6B*.

Immunohistochemistry. Third instar wandering larvae were dissected at room temperature (RT) in Grace's medium (Gibco) and fixed for 10 min in Grace's medium containing 5.3% EM-grade paraformaldehyde (Electron Microscopy Science, PA), freshly used every 10 days. After twice washed for 15 min in PBT (PBS containing 0.1% TritonX-100) and blocked for at least 30 min in PBT, containing 5% normal goat serum (NGS) and 20 μg/mL NaN_3, samples were incubated with primary antibodies in blocking solution overnight. Subsequently, samples were washed 3 times for 20 min in PBT and secondary antibodies were applied in PBS for 4 h at RT. After 3 washes for 20 min with PBT and equilibration in PBS, discs were dissected from the fixed carcasses and mounted in Vectashield (Vector Laboratories). Images were obtained within 24 h using a Leica SP5 Confocal setup. Antibody dilutions used were: rat anti-dCREB-A 1/2000 (Deborah Andrew), rat anti-Dan 1/300 (Steven Cohen), rabbit anti-Salm 1/25 (Reinhard Schuh), guinea pig anti-Hth 1/2000 (Richard Mann), mouse anti-Dll 1/1000 (Steven Cohen), mouse anti-Elav 1/100 (Developmental Studies Hybridoma Bank, University of Iowa), rabbit anti-GFP 1/500 (Molecular Probes), mouse anti-GFP 1/250 (Molecular Probes), rabbit anti-Scr 1/300 (5), and mouse anti-Ey 1/100.

Scanning Electron Microscopy (SEM). Adult flies were fixed for 2 h in 2.5% glutaraldehyde containing 0.1% TritonX-100, washed once for 30 min in PBS, twice for 30 min in water, and passed through increasing dilutions of ethanol in water (30%, 50%, 70%, 90% ethanol in water) for 15 min each time. Finally, cuticles were completely dehydrated in 100% ethanol overnight. After removal of the ethanol and complete air-dying, the heads were dissected and mounted on a SEM block, subjected to gold-coating and observed at the Center for Microscopy of the University of Basel (ZMB) using a Philips XL30 FEG ESEM electron microscope.

High-Resolution APD Imaging. High-resolution APD imaging was performed on a uniquely modified ConfoCor3 instrument (Carl Zeiss, Jena, Germany), consisting of an inverted microscope for transmitted light and epifluorescence (Axiovert 200 M); a VIS-laser module comprising the Ar/ArKr (458, 477, 488 and 514 nm), HeNe 543 nm and HeNe 633 nm lasers; and the scanning module LSM 510 META. The instrument was modified to enable detection using silicon Avalanche Photo Detectors (SPCM-AQR-1X; PerkinElmer, USA) for imaging and FCS. Images were recorded at a 512×512 pixel resolution. The C-Apochromat $40 \times /1.2$ W UV-VIS-IR objective was used throughout.

1. Staehling-Hampton K, Jackson PD, Clark MJ, Brand AH, & Hoffmann FM (1994) Specificity of bone morphogenetic protein-related factors: cell fate and gene expression changes in Drosophila embryos induced by decapentaplegic but not 60A. *Cell Growth Differ* 5(6):585–593.
2. Calleja M, Moreno E, Pelaz S, & Morata G (1996) Visualization of gene expression in living adult Drosophila. *Science* 274(5285):252–255.
3. Grieder NC, Morata G, Affolter M, & Gehring WJ (2009) Spalt major controls the development of the notum and of wing hinge primordia of the Drosophila melanogaster wing imaginal disc. *Dev Biol* 329(2):315–326.
4. Wagner-Bernholz JT, Wilson C, Gibson G, Schuh R, & Gehring WJ (1991) Identification of target genes of the homeotic gene Antennapedia by enhancer detection. *Genes Dev* 5(12B):2467–2480.
5. LeMotte PK, Kuroiwa A, Fessler LI, & Gehring WJ (1989) The homeotic gene Sex combs reduced of Drosophila: Gene structure and embryonic expression. *EMBO J* 8(1):219–227.
6. Zhao JJ, Lazzarini RA, & Pick L (1996) Functional dissection of the mouse Hox-a5 gene. *EMBO J* 15(6):1313–1322.
7. Tour E, Hittinger CT, & McGinnis W (2005) Evolutionarily conserved domains required for activation and repression functions of the Drosophila Hox protein Ultrabithorax. *Development* 132(23):5271–5281.

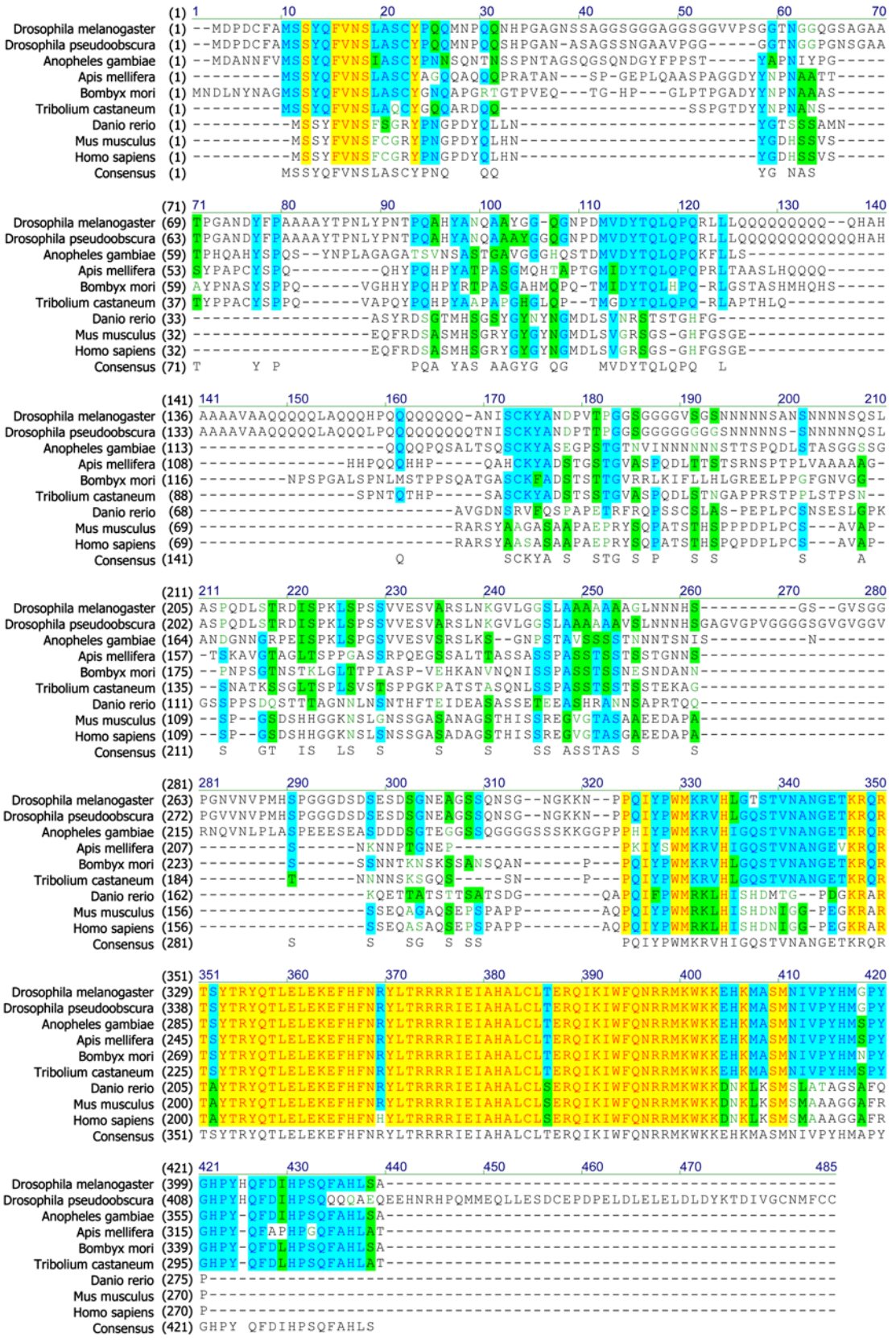

Fig. S1. Alignment of Scr amino acid sequence in insects and vertebrates. The YPWM motif, the homeodomain and the N-terminal MSSYFVNS motif are conserved among all animals examined, yet only the first two play a role in transcriptional specificity of Scr in vivo (see text for more details). The MSSYFVNS motif participates in enhancing transcriptional potency in the mouse Scr homolog Hoxa5 (6). The DYTQL motif, also described to enhance transcriptional activation (7), is only conserved among insect species. Little is known about the function of the SCKYA motif. Note the high conservation of the linker between HD and YPWM motif in insects. The same applies to the C-terminal sequences of Scr.

Quantitative study of synthetic Hox transcription factor–DNA interactions in live cells

Vladana Vukojević[a,1], Dimitrios K. Papadopoulos[b,1], Lars Terenius[a], Walter J. Gehring[b,2], and Rudolf Rigler[c,d,2]

[a]Department of Clinical Neuroscience, Karolinska Institutet, 17176 Stockholm, Sweden; [b]Department of Cell Biology, Biozentrum, University of Basel, 4056 Basel, Switzerland; [c]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 17177 Stockholm, Sweden; and [d]Laboratory of Biomedical Optics, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland

Contributed by Walter J. Gehring, December 23, 2009 (sent for review November 3, 2009)

Transcription factor–DNA interactions are life sustaining and therefore the subject of intensive research. In spite of vast effort, quantitative in vivo studies of the molecular mechanisms underlying these fundamental interactions remain challenging. In the preceding paper, we designed synthetic Sex combs reduced (Scr) peptides and validated genetically their function as transcriptional regulators. Here we present a controllable system for quantitative studies of protein–DNA interactions in live cells that enables us to "titrate" the concentration of the synthetic Scr peptides in a single cell. Using methods with single-molecule sensitivity, advanced fluorescence imaging and fluorescence correlation spectroscopy (FCS), we were able to study the kinetics of Scr-DNA interactions in live salivary gland cells, where *Scr* is normally expressed during development. We discerned freely moving Scr molecules, characterized the specific and nonspecific Scr peptide–DNA interactions, and estimated their corresponding dissociation constants (K_d) in vivo. Our results suggest that the synthetic Scr transcription factors find their specific target sites primarily by multiple association/dissociation events, the rapidity of which is largely owed to electrostatic interactions. Based on these new findings, we formulate a model mechanism and emulate the kinetics of Scr homeodomain–DNA interactions in live cells using numerical simulations.

fluorescence correlation spectroscopy | single-molecule sensitivity | homeodomain | Sex combs reduced | synthetic peptides

Transcription factor–DNA interactions, the assembly of functional transcriptional complexes, and the mechanisms involved in target site recognition are central to understanding gene regulation in vivo. In spite of such relevance, kinetic studies of transcription factor–DNA interactions are limited and have been performed only for a handful of proteins. Most prominent studies include the lac repressor in *Escherichia coli* (1–4), nuclear receptor-directed transcription (5), helicase translocation along DNA (6), and restriction enzymes–DNA interactions (7, 8). The majority of these studies are carried out with naked DNA or reconstituted chromatin, reflecting the in vivo situation only to a limited extent. Chromatin structure in live cells is highly dynamic, involving multiple interactions and transformations (9–11). Moreover, a number of cofactors and coactivators/corepressors are involved in transcriptional regulation, suggesting that gene expression is accomplished through a dynamic, spatiotemporally entangled interplay between concomitant processes—protein searching for specific target sites, rearrangements of DNA conformation, and recruitment of cofactors and/or coregulators (5, 12, 13). Such complexity is not easily mimicked in solution, therefore necessitating live cell experimentation. However, these measurements are complex, and kinetic studies of protein–DNA interactions in live cells remain scarce (3, 4, 14, 15), mainly due to the limited number of experimental approaches with single-molecule sensitivity that enable nondestructive observation of molecular interactions in live cells (16–18). Recently, the visualization of transcription at native loci has been performed in live cells using two-photon excitation microscopy (19–21). This work was done by gene overexpression, which is an experimental artifact that needs to be overcome to enable quantitative studies of protein–DNA interactions in live cells.

Here we study quantitatively the molecular mechanisms of homeodomain (HD) – DNA interactions in live cells, which is the first step in Hox-mediated transcription. Using functional and nonfunctional synthetic variants of the *Drosophila Hox* gene *Sex combs reduced* (*Scr*), the activity of which has been studied in vivo by phenotypic and genetic analyzes (22), we exploit the minimal expression levels of the *hsp70* promoter of the UAS constructs to express the transcription factor in concentrations as close as possible to native conditions. Using confocal laser scanning microscopy with avalanche photodiodes (APDs), so-called APD imaging (23), and fluorescence correlation spectroscopy (FCS) (24–26) we have established a high-resolution experimental modality that enables nondestructive observation of transcription factor molecules and quantitative studies of molecular interactions in live cells with single-molecule sensitivity. We use APD imaging to visualize the Scr-HD molecules at low expression levels and FCS to study quantitatively Scr-HD numbers and mobility. This is done by monitoring the fluorescence intensity fluctuations in a selected spot in the salivary gland nuclei that is generated by focusing the incident laser light through the microscope objective. Statistical analysis of the recorded data is applied in order to derive molecular numbers and macroscopic diffusion constants. For recent reviews on FCS, see, for example refs. 27–29. A short description of the methodology is given in *SI Text*.

Results

Controllable Hox Peptide Expression in Live Cells. In order to study the interactions of synthetic HD peptides with nuclear DNA in live cells we first composed a controllable live cell expression system with properties as close as possible to native conditions. We used synthetic *Scr* genes in live salivary gland cells, where *Scr* is normally expressed during development (30). Substitutions of threonine 6 and serine 7 of the Scr-HD by alanines (Scr-HD_{AA}) or aspartates (Scr-HD_{DD}) render Scr constitutively active or inactive, respectively (31). Additional mutations of residues 50 and 51 of the third helix of the HD into alanines (Scr-$HD_{DD}^{Q50AN51A}$) completely abolish binding of the HD to the DNA (32). These active and inactive variants allowed us to analyze and compare the differential behavior of functional transcription factors, as compared to their nonfunctional counterparts.

To select a fluorescent probe with optimal photophysical properties for FCS analysis, we used the Gal4-UAS system

Author contributions: V.V., D.K.P., W.J.G., and R.R. designed research; V.V. and D.K.P. performed research; V.V., D.K.P., L.T., and R.R. contributed new reagents/analytic tools; V.V., D.K.P., and R.R. analyzed data; and V.V., D.K.P., L.T., W.J.G., and R.R. wrote the paper.

The authors declare no conflict of interest.

[1]These authors contributed equally to this work

[2]To whom correspondence should be addressed. E-mail: walter.gehring@unibas.ch or rudolf.rigler@ki.se.

This article contains supporting information online at www.pnas.org/cgi/content/full/0914612107/DCSupplemental.

and generated UAS lines encoding the synthetic Scr peptides fused to different fluorescent proteins (Fig. 1*A*). Under our experimental conditions the mCitrine–Scr-HD fusions were characterized by the highest photon count per molecule (2–5 kHz). We therefore selected this fluorescence protein as the most appropriate for FCS analysis and used it throughout the study.

We also evaluated different expression systems (Fig. 1). The use of salivary gland-specific Gal4 drivers (*sgs3*-Gal4 and *dpp*^{blink}-Gal4) resulted in *Scr*-HD overexpression, with concentrations in nuclei often exceeding 1 μM (300 nM – 3 μM) (Fig. 1*B* and *C*). At such high expression, experimental artifacts are introduced—the transcription factor is distributed all over the salivary gland nuclei (Fig. 1*B*), and the nonspecific DNA–Scr-HD interactions are pronounced, notably concealing the specific interactions (Fig. 1*C*). Therefore, low expression systems are required. Taking advantage of expression "leakage" of the *hsp70* minimal promoter of the UAS constructs (Fig. 1*A*), we were able to express the *mCitrine-Scr-HD* variants without the use of any Gal4 driver (Fig 1*D*) and regulate the expression level of Scr peptides by heat shock (Fig. 2). Heat shock triggers a physiological response known to increase the transcription levels of the *hsp* genes, while decreasing the expression levels of most other genes (33). To ensure the expression of all constructs, immunostaining of salivary glands using *yw*^{1118} as a negative control (injection background) was performed. All lines exhibited substantial nuclear staining of Scr-HD as compared to the control (Fig. S1).

APD imaging of polytene nuclei and FCS confirmed that the expression levels of the synthetic Scr-HD peptides increased gradually after heat shock (Fig. 2*A*–*D*). In addition, FCS analysis revealed marked differences in Scr-HD dynamics depending on its concentration in the same cell (Fig. 2*D* and *E*). At low Scr-HD concentrations, sigmoid shaped autocorrelation curves with only one characteristic time (τ_1) were observed (Fig. 2*D* and *E*, *Blue Curve*). As the Scr-HD concentration increases, nonspecific interactions with nuclear DNA become prominent. This is reflected in FCS measurements by changes in the shape of the autocorrelation curve. For increasing concentration the autocorrelation curves become complex and a second component with a distinctly longer characteristic time (τ_2) was observed (Fig. 2*D* and *E*). The relative amount of the second component, which is reflected by its corresponding amplitude, increases for increasing Scr-HD concentrations (Fig. 2*F*).

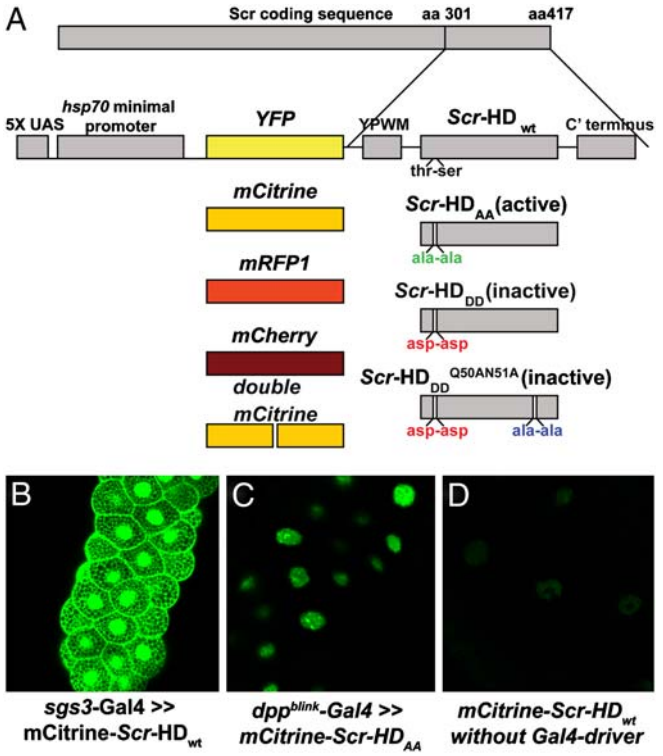

Fig. 1. UAS lines generated and quantitative comparison between Gal4-mediated overexpression and *hsp70* basic expression levels by APD imaging. (*A*) Fluorescent protein variants (*YFP*, *mCitrine*, *mRFP1*, *mCherry* and double *mCitrine*) were used to generate 5′ fusions of the four synthetic *Scr*-HD genes and to generate transgenic flies. The Scr-HD synthetic peptides are comprised of the terminal 117 amino acids of the 417 amino acids long Scr protein. Constructs are not drawn to scale. (*B*) The salivary gland-specific driver *sgs3*-Gal4 expressed *mCitrine-Scr-HD*$_{wt}$ at very high levels, inducing an artificial distribution of the transcription factor all over the salivary gland cells. (*C*) *dpp*^{blink}-Gal4 resulted in significantly weaker expression levels (1–3 μM) and correct nuclear localization but was still difficult to study by FCS because of the high expression levels. (*D*) Minimal expression levels exhibited by the UAS constructs (50–300 nM), without the use of any Gal4 driver, facilitated the study of molecular movement and interactions by FCS.

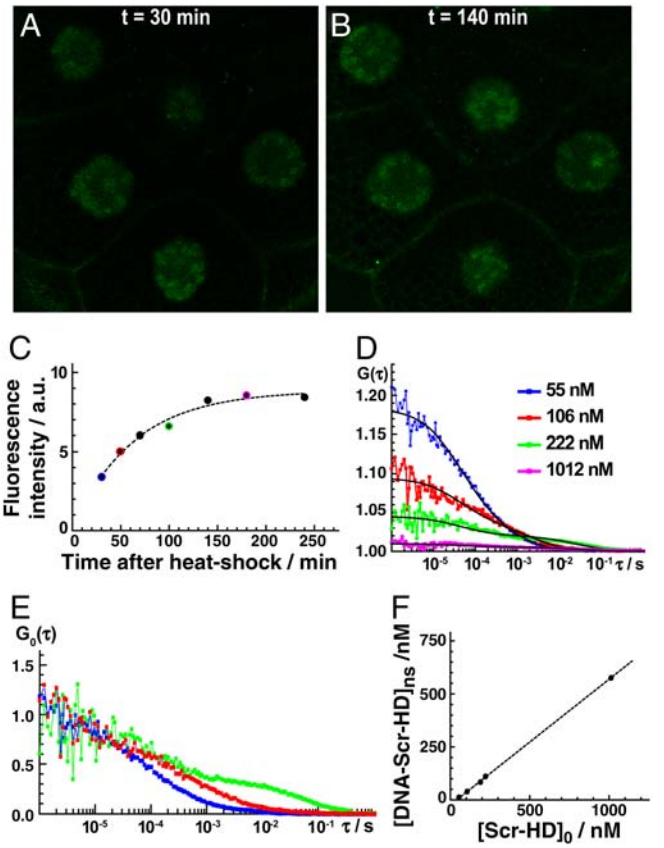

Fig. 2. Controllable *Scr*-HD expression in vivo. (*A*) APD image of salivary gland nuclei expressing the synthetic Scr-HD$_{wt}$ transcription factor. Heat shock was applied by keeping the *Drosophila* larvae at 37 °C for 30 min. The image was recorded 30 min after heat shock, when the expression was still low and close to physiological levels. Minimal expression was obtained by the UAS construct without the Gal4 driver. (*B*) APD imaging showing elevated expression of *Scr*-HD$_{wt}$ 140 min after heat shock. (*C*) Gradual increase of Scr-HD$_{wt}$ concentration in live salivary gland cells recorded by time-lapse APD imaging. (*D*) Gradual increase of Scr-HD$_{wt}$ concentration in live salivary gland cells recorded by FCS. The color code indicates FCS measurements performed just after APD imaging, represented in (*C*) by the corresponding color. (*E*) Normalized autocorrelation curves showing changes in Scr-HD$_{wt}$ dynamics as a function of its concentration. All FCS measurements were performed at the same location in the same cell. (*F*) The concentration of the second component increases linearly with the total concentration of Scr-HD$_{wt}$.

Scr-HD Variants Show Differences in Interaction with DNA at the Molecular Level. At low concentrations of the Scr-HD transcription factors, 30 nM < [Scr-HD] < 200 nM, we could now clearly visualize marked differences in molecular distribution and dynamics among Scr-HD_{wt}, Scr-HD_{AA}, Scr-HD_{DD}, and Scr-$HD_{DD}^{Q50AN51A}$ (Fig. 3). While the wild type (Fig. 3*A*) and the constitutively active Scr-HD_{AA} (Fig. 3*B*) variants showed patterned distribution, staunchly reflecting the contours of the polytene chromosomes, Scr-HD_{DD} was uniformly distributed in the nucleus (Fig. 3*C*), and Scr-$HD_{DD}^{Q50AN51A}$ was excluded from chromatin and appeared to be "squeezed" between polytene chromosomes (Fig. 3*D*). FCS analysis revealed that the Scr-HD variants differ in their ability to interact with nuclear DNA. These differences that are readily observed at low Scr-HD concentrations (Fig. 3*E*) become less discernible at high Scr-HD concentrations (Fig. 3*F*).

In order to identify processes that generate the fluorescence intensity fluctuations, we examined the effect of the detection volume size on the characteristic times τ_1 and τ_2 (Fig. 4). Both characteristic times increase when the detection volume element is enlarged (Fig. 4*C*), indicating that the fluorescence intensity fluctuations are generated by molecular movement. Linear dependence of the characteristic times on the median detection volume element area (Fig. 4*D* and *E*) suggests that Scr-HD molecules move in a free diffusion-like fashion, showing no indication of anomalous diffusion.

FCS analysis revealed also that the dynamics of Scr-HD peptides is different in different regions of the nucleus. In the nucleoplasm, where DNA concentration is generally low, Scr-HD movement is faster than on the polytene chromosomes and the sites of pronounced Scr-HD accumulation (Fig. S2).

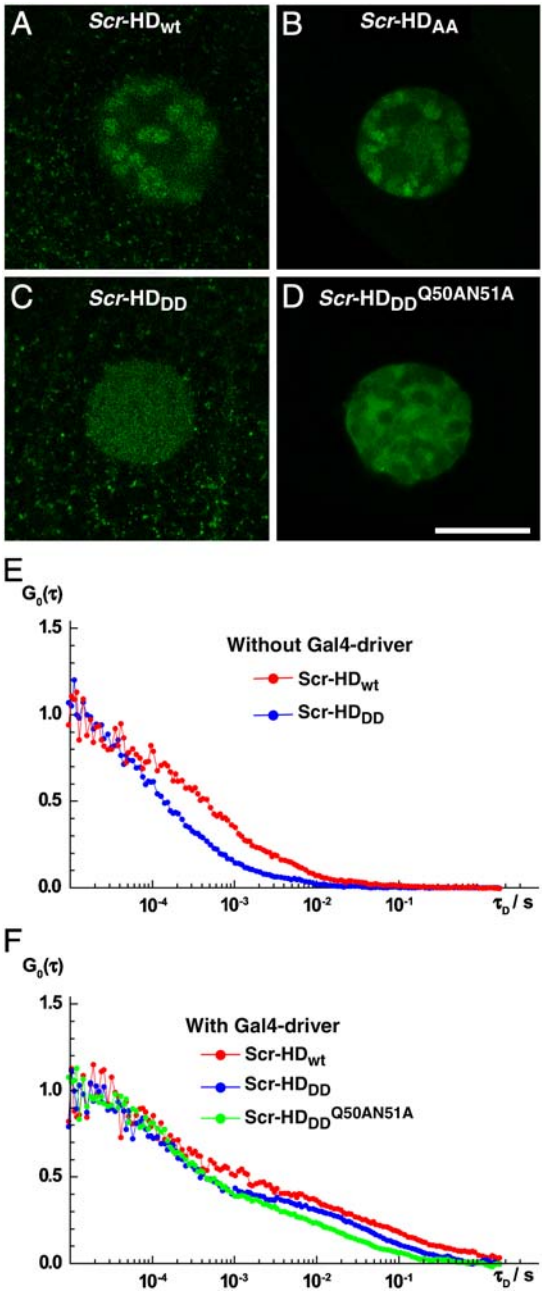

Fig. 3. Differences in DNA–Scr-HD interactions between different Scr-HD variants visualized in polytene nuclei of live salivary gland cells by APD imaging and FCS. (*A–D*) Polytene nuclei expressing the synthetic *Scr*-HD_{wt}, *Scr*-HD_{AA}, *Scr*-HD_{DD} and *Scr*-$HD_{DD}^{Q50AN51A}$ at low, physiologically relevant concentrations [Scr-HD_{wt}] = 105 nM, [Scr-HD_{AA}] = 157 nM, [Scr-HD_{DD}] = 66 nM and [Scr-$HD_{DD}^{Q50AN51A}$] = 220 nM. (*E–F*) Differences in DNA–Scr-HD interactions between the transcriptionally active and inactive variants can be readily observed at low but are less distinguishable at high Scr-HD concentrations.

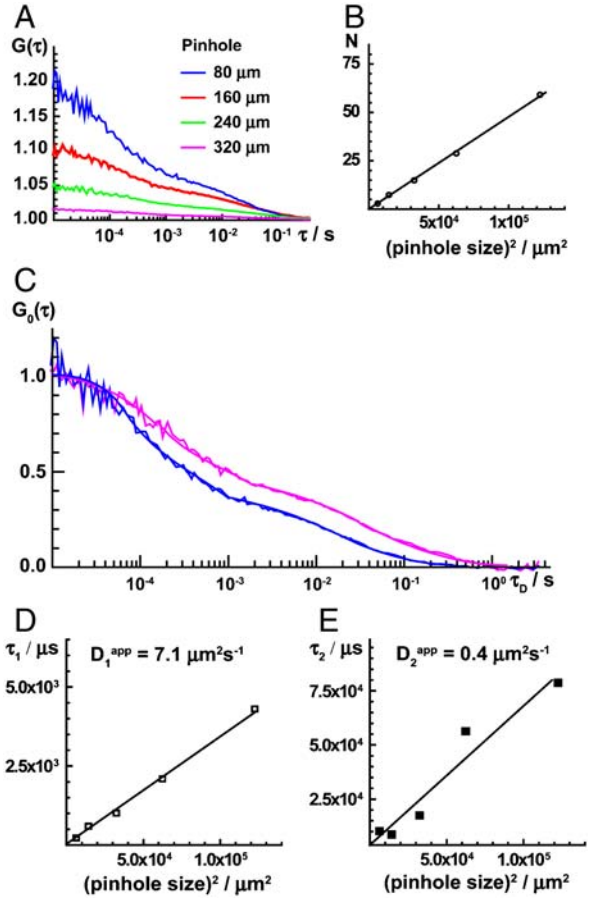

Fig. 4. Fluorescence intensity fluctuations recorded in live salivary gland nuclei are generated by molecular movement. (*A*) Autocorrelation curves recorded in detection volume elements of different size. The amplitude of the autocorrelation curve decreases, indicating that the number of observed Scr-HD molecules increases as the detection volume element is enlarged. (*B*) The number of Scr-HD molecules in the detection volume element increases linearly with the median surface area of the detection volume element. (*C*) Normalized autocorrelation curves showing that both characteristic times, τ_1 and τ_2, increase when the detection volume element is enlarged. The same color code as in (*A*) is used. (*E* and *F*) Both characteristic times, τ_1 (*E*) and τ_2 (*F*), increase linearly with the median surface area of the detection volume element, suggesting that Scr-HD molecules move in a free diffusion-like fashion.

Discerning Specific from Nonspecific DNA–Scr-HD Interactions in Live Cells and Determination of DNA–Scr-HD Binding Constants from FCS Measurements. In order to construct DNA–Scr-HD binding curves (Fig. 5), we exploited the responsiveness of the *hsp70* minimal promoter of the UAS constructs to heat shock to generate different *Scr*-HD express levels. Autocorrelation curves recorded in individual nuclei were analyzed using a two-component model with triplet formation (25, 28). FCS analysis revealed that Scr-HD_{DD} and Scr-$HD_{DD}^{Q50AN51A}$ binding increased linearly with their concentration in a fashion that is characteristic for nonspecific binding (Fig. 5, *Black Line*). In contrast, Scr-HD_{wt} and Scr-HD_{AA} binding profiles showed a "shifted" line, suggesting that interactions other than the nonspecific ones are also involved (Fig. 5, *Red Line*). The same shifted binding profile was observed in individual cells (Fig. 2*F*) and in a collection of cells (Fig. 5*B*, *Red Line*). In order to understand these results, we proceeded to analyze the kinetics of the underlying chemical interactions numerically.

Dissociation Constant of the Nonspecific DNA – Scr-HD_{DD} and DNA-Scr-$HD_{DD}^{Q50AN51A}$ Complexes. Previous genetic experiments (22), imaging analysis, and FCS measurements suggest that Scr-HD_{DD} and Scr-$HD_{DD}^{Q50AN51A}$ interact nonspecifically with the DNA. These interactions, exemplified for Scr-HD_{DD}, can be described as a reversible chemical reaction:

$$\mathrm{DNA} + \mathrm{ScrHD_{DD}} \underset{k_{-ns}}{\overset{k_{ns}}{\Leftrightarrow}} \mathrm{DNA - ScrHD_{DD}}, \qquad [1]$$

where k_{ns} and k_{-ns} are rate constants for the formation and dissociation of the nonspecific DNA – Scr-HD_{DD} complex, respectively.

Using mass balance equations (*SI Text*), we could show that the concentration of nonspecific DNA – Scr-HD_{DD} complex increases linearly with the total concentration of the transcription factor:

$$[\mathrm{DNA - ScrHD_{DD}}] = \frac{k_{ns}\cdot[\mathrm{DNA}]_0}{(k_{-ns}+k_{ns}\cdot[\mathrm{DNA}]_0)}\cdot[\mathrm{ScrHD_{DD}}]_0. \qquad [2]$$

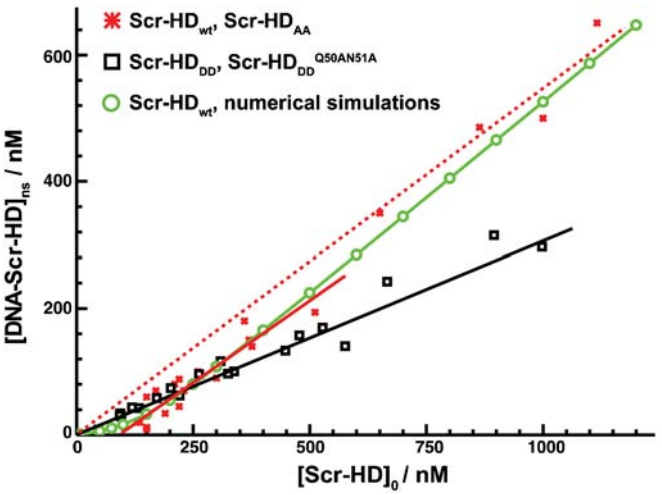

Fig. 5. In vivo DNA – Scr-HD binding. FCS measurements show that the concentration of DNA – Scr-HD_{DD} and DNA – Scr-$HD_{DD}^{Q50AN51A}$ complexes increases linearly with the total concentration of the transcription factor (*Black Line*), suggesting, in accordance with Eq. **2**, that Scr-HD_{DD} and Scr-$HD_{DD}^{Q50AN51A}$ interact only nonspecifically with the genomic DNA. In contrast, Scr-HD_{AA} and Scr-HD_{wt} undergo both specific and nonspecific interactions with the DNA (*Solid Red Line*), showing a binding profile that is in accordance with Eq. **7**. In the absence of specific DNA – Scr-HD_{wt} interactions, the nonspecific DNA – Scr-HD_{wt} binding would follow the dashed red line. DNA – Scr-HD_{wt} interactions were also numerically simulated (*Green Curve*) as described in the text.

Using Eq. **2** and the experimentally determined slope (Fig. 5, *Black Line*)

$$\frac{k_{ns}\cdot[\mathrm{DNA}]_0}{(k_{-ns}+k_{ns}\cdot[\mathrm{DNA}]_0)} = (0.35\pm0.15), \qquad [3]$$

we derived the equilibrium dissociation constant for the nonspecific DNA – Scr-HD_{DD} complex:

$$\frac{k_{-ns}}{k_{ns}} = K_d^{\mathrm{ScrHD_{DD,ns}}} = \frac{[\mathrm{DNA}]_0}{0.54}, \qquad [4]$$

where $[\mathrm{DNA}]_0$ is the total concentration of potential sites for nonspecific binding.

The total concentration of potential binding sites for nonspecific interactions is not easily estimated. Following the assumptions detailed out in *SI Text*, we estimated that $[\mathrm{DNA}]_0 =$ 46 μM. Under these assumptions, the dissociation constant for nonspecific DNA–Scr-HD binding is estimated to be:

$$K_d^{\mathrm{ScrHD_{DD,ns}}} = (80\pm50)\ \mu\mathrm{M}. \qquad [5]$$

Dissociation Constant of the Specific DNA – Scr-HD_{wt} and DNA – Scr-HD_{AA} Complexes. The transcriptionally active variants Scr-HD_{wt} and Scr-HD_{AA} undergo both specific and nonspecific interactions with the DNA. Supposing that nonspecific interactions precede the specific ones, a two-step process of consecutive reactions was assumed:

$$\mathrm{DNA_{ns}} + \mathrm{ScrHD_{wt}} \underset{k_{-ns}}{\overset{k_{ns}}{\Leftrightarrow}} (\mathrm{DNA - ScrHD_{wt}})_{ns}$$
$$(\mathrm{DNA - ScrHD_{wt}})_{ns} + \mathrm{DNA_s} \underset{k_{-s}}{\overset{k_s}{\Leftrightarrow}} (\mathrm{DNA - ScrHD_{wt}})_s. \qquad [6]$$

Applying the quasi-steady state approximation (*SI Text*), we could show that:

$$[(\mathrm{DNA - ScrHD_{wt}})_{ns}] = \frac{k_{ns}\cdot[\mathrm{DNA}]_0}{k_{-ns}+k_s\cdot[\mathrm{DNA_s}]+k_{ns}\cdot[\mathrm{DNA}]_0}\cdot[\mathrm{ScrHD_{wt}}]_0 - \frac{k_{ns}\cdot[\mathrm{DNA}]_0 - k_{-s}}{k_{-ns}+k_s\cdot[\mathrm{DNA_s}]+k_{ns}\cdot[\mathrm{DNA}]_0}\cdot[(\mathrm{DNA - ScrHD_{wt}})_s], \qquad [7]$$

in agreement with our experimental findings by FCS (Fig. 2*F* and Fig. 5, *Solid Red Line*). The slope gives:

$$\frac{k_{ns}\cdot[\mathrm{DNA}]_0}{k_{-ns}+k_s\cdot[\mathrm{DNA_s}]+k_{ns}\cdot[\mathrm{DNA}]_0} = (0.60\pm0.05) \qquad [8]$$

and the intercept:

$$\frac{k_{ns}\cdot[\mathrm{DNA}]_0 - k_{-s}}{k_{-ns}+k_s\cdot[\mathrm{DNA_s}]+k_{ns}\cdot[\mathrm{DNA}]_0}\cdot[(\mathrm{DNA - ScrHD_{wt}})_s] = (50\pm30)\ \mathrm{nM}. \qquad [9]$$

If k_{-s} is small compared to $k_{ns}\cdot[\mathrm{DNA}]_0$ and can therefore be neglected, then

$$0.6\cdot[(\mathrm{DNA - ScrHD_{wt}})_s] = 50\ \mathrm{nM}, \qquad [10]$$

and the total concentration of the specific complex DNA–Scr-HD_{wt} in the polytene nucleus could be estimated

$$[(\mathrm{DNA - ScrHD_{wt}})_s] = (80\pm50)\ \mathrm{nM}. \qquad [11]$$

Using the estimated concentration of specific DNA – Scr-HD$_{wt}$ complexes [Eq. **11**] and the total concentration of specific DNA binding sites [DNA$_s$] = 92 nM (*SI Text*), the dissociation constant for the specific DNA – Scr-HD$_{wt}$ complex was estimated:

$$K_d^{\mathrm{ScrHD_{wt,s}}} = \frac{[\mathrm{DNA}_s]_{\mathrm{free}} \cdot [\mathrm{ScrHD_{wt}}]_{\mathrm{free}}}{[(\mathrm{DNA-ScrHD_{wt}})_s]} = \frac{12\ \mathrm{nM} \cdot 48\ \mathrm{nM}}{80\ \mathrm{nM}} = (7 \pm 5)\ \mathrm{nM}. \quad [12]$$

Dissociation Constant of the Nonspecific DNA – Scr-HD$_{wt}$ and DNA – Scr-HD$_{AA}$ Complexes. FCS data enabled us also to estimate the dissociation constant for the nonspecific DNA – Scr-HD$_{wt}$ and DNA – Scr-HD$_{AA}$ complexes. Using the analysis presented above and the data from Fig. 6 (*Dotted Red Line*) we could estimate the dissociation constant for the nonspecific DNA – Scr-HD$_{wt}$ and DNA – Scr-HD$_{AA}$ complexes

$$\frac{k_{\mathrm{ns}} \cdot [\mathrm{DNA}]_0}{(k_{-\mathrm{ns}} + k_{\mathrm{ns}} \cdot [\mathrm{DNA}]_0)} = (0.65 \pm 0.05) \quad [13]$$

$$\frac{k_{-\mathrm{ns}}}{k_{\mathrm{ns}}} = K_d^{\mathrm{ScrHD_{wt,ns}}} = \frac{[\mathrm{DNA}]_0}{1.9}. \quad [14]$$

Under the assumptions outlined in *SI Text*, we estimated

$$\frac{k_{-\mathrm{ns}}}{k_{\mathrm{ns}}} = K_d^{\mathrm{ScrHD_{wt,ns}}} = (25 \pm 15)\ \mu\mathrm{M}. \quad [15]$$

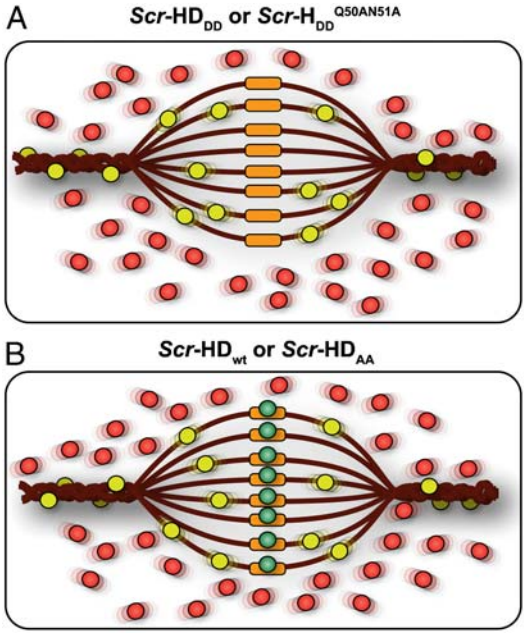

Fig. 6. Schematic representation of the differential behavior of Scr-HD variants. (*A*) In nuclei expressing the transcriptionally inactive variants *Scr*-HD$_{DD}$ or Scr-HD$_{DD}^{Q50AN51A}$ FCS can distinguish between free Scr-HD molecules diffusing in the nucleoplasm (*Red Spheres*) and transcription factor molecules interacting nonspecifically with the DNA (*Yellow Spheres*). Binding sites within the region of loose chromatin conformation (*Orange Rectangles*) cannot be bound specifically by Scr-HD$_{DD}$ or Scr-HD$_{DD}^{Q50AN51A}$ because binding is abolished by the corresponding amino acid substitutions. (*B*) DNA – Scr-HD$_{wt}$ or DNA – Scr-HD$_{AA}$ interactions are more complex. In addition to the free Scr-HD molecules in the nucleoplasm (*Red Spheres*) and transcription factor molecules interacting nonspecifically with the DNA (*Yellow Spheres*), the transcriptionally active variants undergo also specific interactions (*Green Spheres*) with putative specific binding sites.

This threefold difference in nonspecific interactions between DNA – Scr-HD$_{DD}$ and DNA – Scr-HD$_{DD}^{Q50AN51A}$ on one side and DNA – Scr-HD$_{wt}$ and DNA – Scr-HD$_{AA}$ on the other can be naturally explained by differences in the intensity of electrostatic interactions with the negatively charged DNA (*SI Text*). Charge analysis revealed that at relevant pH (34) Scr-HD$_{wt}$ and Scr-HD$_{AA}$ are more positively charged than Scr-HD$_{DD}$ and Scr-HD$_{DD}^{Q50AN51A}$, explaining also the difference in their mobility observed by FCS (Fig. 3*E*).

Numerical Simulation of DNA – Scr-HD$_{wt}$ Interactions. FCS measurements provided important, experimentally derived relationships between rate constants, enabling us to take one step further and attempt to simulate the dynamics of DNA – Scr-HD interactions. Using the experimentally determined relations between rate constants given in Eqs. **3**, **8**, and **9** and the estimated concentration of specific and nonspecific binding sites we derived by fitting the macroscopic rate constant for Scr-HD$_{wt}$ interactions with the DNA, $k_{\mathrm{ns}} = 1 \times 10^7\ \mathrm{M^{-1}s^{-1}}$, $k_{-\mathrm{ns}} = 3.25 \times 10^2\ \mathrm{s^{-1}}$, $k_s = 1.5 \times 10^7\ \mathrm{M^{-1}s^{-1}}$, and $k_{-s} = 1 \times 10^{-1}\ \mathrm{s^{-1}}$. Using these values, the experimental data derived by FCS could be readily fitted by the simple two-step model given in **6** (Fig. 5, *Green Curve*).

Discussion

Our results suggest that the synthetic Scr-HD transcription factors find their specific target sites primarily by multiple association/dissociation events, the rapidity of which is largely owed to electrostatic interactions. A model mechanism for DNA – Scr-HD interactions is schematically represented in Fig. 6. By advanced APD imaging and FCS, we detected free Scr-HD molecules that move in the nucleoplasm at a rate that is similar to free 3D diffusion of the purified recombinant mCitrine-Scr-HD$_{wt}$ peptide in solution, $D_{\mathrm{ScrHD_{wtsolution}}} = (2 \pm 1) \times 10^{-11}\ \mathrm{m^2s^{-1}}$ (Fig. 6, *Red Spheres*). On the chromosomes and in regions of pronounced accumulation, the Scr-HD molecules undergo numerous nonspecific interactions with the DNA (Fig. 6, *Yellow Spheres*). The dissociation constant for the nonspecific DNA – Scr-HD complexes was estimated to be $K_d^{\mathrm{ScrHD_{DD,ns}}} = (80 \pm 50)\ \mu\mathrm{M}$ for DNA – Scr-HD$_{DD}$ and Scr-HD$_{DD}^{Q50AN51A}$ and $K_d^{\mathrm{ScrHD_{wt,ns}}} = (25 \pm 15)\ \mu\mathrm{M}$ for DNA–Scr-HD$_{wt}$ and DNA – Scr-HD$_{AA}$. These somewhat high K_d values suggest that Scr-HD affinity for nonspecific binding sites is rather low, a feature that allows the DNA–Scr-HD complex to assemble/disassemble rapidly in the polytene nucleus densely packed with DNA. Scr-HD molecules undergoing nonspecific interactions with the nuclear DNA are slowed down by the on/off interactions and seem to move in a diffusion-like fashion with an apparent diffusion constant $D = (4 \pm 2) \times 10^{-13}\ \mathrm{m^2s^{-1}}$.

Nonspecific DNA – Scr-HD$_{wt}$ interactions always precede the specific ones. However, if the binding site is a putative specific binding site in a region of loose chromatin conformation (Fig. 6, *Orange Rectangles*), the nonspecific complex readily transforms into a specific complex (Fig. 6, *Green Spheres*). The number of specific binding sites is lower than the number of nonspecific sites, and the mobility of the specific DNA–Scr-HD complexes seems to be rather low, rendering these complexes "invisible" for direct observation by FCS. Nevertheless, we were able to characterize the specific interactions using rigorous kinetic analysis that was based on FCS measurements performed at very low Scr-HD concentrations. We estimated the dissociation constant of the specific DNA – Scr-HD$_{wt}$ complexes to be in the nanomolar range, $K_d^{\mathrm{ScrHD_{wt,s}}} = (7 \pm 5)\ \mathrm{nM}$, and their concentration $[(\mathrm{DNA-ScrHD_{wt}})_s] = (80 \pm 50)\ \mathrm{nM}$.

The single-molecule sensitivity of APD imaging and FCS enabled us to visualize and quantitatively characterize subtle differences between molecular interactions in live cells, a result hitherto not achieved by any other method of analysis. These

results enabled us to probe the kinetics of Scr-HD binding to nuclear DNA using numerical simulations. We derived by fitting macroscopic rate constants for formation/dissociation of the DNA – Scr-HD complexes: $k_{ns} = 1 \times 10^7$ $M^{-1}s^{-1}$ ($k_{ns} = 4.5\times 10^6$ $M^{-1}s^{-1}$ for Scr-HD_{DD} and Scr-$HD_{DD}^{Q50AN51A}$), $k_{-ns} = 3.25\times 10^2$ s^{-1}, $k_s = 1.5 \times 10^7$ $M^{-1}s^{-1}$, and $k_{-s} = 1 \times 10^{-1}$ s^{-1} and emulated the DNA – Scr-HD interactions by numerical simulations. The rate constant for specific DNA – Scr-HD_{wt} interactions derived by fitting, $k_s = 1.5 \times 10^7$ $M^{-1}s^{-1}$, is in agreement with measurements of protein association rate constants to a specific site in a long DNA molecule showing that most proteins bind to their specific sites with an association rate constant of the order 1×10^7 $M^{-1}s^{-1}$ (35).

This study has not only facilitated our understanding of Hox-mediated gene regulation at the molecular level, but also defined an experimental modality for quantitative and nondestructive study of protein–DNA interactions in live cells.

Material and Methods

Preparation of Salivary Glands for FCS and Heat-Shock Procedures. For FCS measurements salivary glands were dissected and transferred to chambered slides (Nalge Nunc International) in PBS. For titration of Scr-HD expression variable heat shocks (10 min to 2 h) were applied to larvae prior to dissection (refer to *SI Methods* for more details).

APD Imaging and FCS. High-resolution fluorescence imaging and FCS measurements were performed on a uniquely modified ConfoCor3 instrument (Carl Zeiss, Jena, Germany) consisting of an inverted microscope for transmitted light and epifluorescence (Axiovert 200 M); a VIS-laser module comprising the Ar/ArKr (458, 477, 488 and 514 nm), HeNe 543 nm and HeNe 633 nm lasers; and the scanning module LSM 510 META. The instrument was modified to enable detection using silicon Avalanche Photo Detectors (SPCM-AQR-1X; PerkinElmer, USA) for imaging and FCS. Images were recorded at a 512×512 pixel resolution. The C-Apochromat $40\times/1.2$ W UV-VIS-IR objective was used throughout. Fluorescence intensity fluctuations were recorded in arrays of 10–30 consecutive measurements, each measurement lasting 5–10 s. Averaged curves were analyzed using the software for online data analysis or exported and fitted offline using the OriginPro 8 data analysis software (OriginLab Corporation, Northampton, MA). In either case, the nonlinear least-square fitting of the autocorrelation curve was performed using the Levenberg–Marquardt algorithm. Quality of the fitting was evaluated by visual inspection and by residuals analysis. The variation between independent measurements reflects variations between cells, rather than imprecision of FCS measurements. For more details on APD imaging and FCS refer to *SI Text*.

Numerical Simulations. Numerical simulations were carried out using the Gepasi 3.30 software package for modeling biochemical systems (36) and software for numerical integration of kinetic equations with the fifth order Runge Kutta method (37).

ACKNOWLEDGMENTS. This work was supported by the Kantons of Basel–Stadt and Basel–Landschaft, a grant from the Swiss National Science Foundation, the European Network of Excellence "Cells into Organs," the Swedish Brain Foundation, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, and the Ministry of Sciences and Technological Development of Serbia (Grants no. 142025 and 142019).

1. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3:318–356.
2. Riggs AD, Bourgeois S, Cohn M (1970) The lac repressor-operator interaction 3. Kinetic studies. *J Mol Biol* 53(3):401–417.
3. Elf J, Li GW, Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science* 316(5828):1191–1194.
4. Wang X, Reyes-Lamothe R, Sherratt DJ (2008) Visualizing genetic loci and molecular machines in living bacteria. *Biochem Soc Trans* 36(Pt 4):749–753.
5. Metivier R, Reid G, Gannon F (2006) Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Rep* 7(2):161–167.
6. Myong S, Rasnik I, Joo C, Lohman TM, Ha T (2005) Repetitive shuttling of a motor protein on DNA. *Nature* 437(7063):1321–1325.
7. Halford SE, Marko JF (2004) How do site-specific DNA-binding proteins find their targets? . *Nucleic Acids Res* 32(10):3040–3052.
8. Szczelkun MD (2000) How do proteins move along DNA? Lessons from type-I and type-III restriction endonucleases. *Essays Biochem* 35:131–143.
9. Cremer T, et al. (2000) Chromosome territories, interchromatin domain compartment, and nuclear matrix: An integrated view of the functional nuclear architecture. *Crit Rev Eukar Gene Expr* 10(2):179–212.
10. Schneider R, Grosschedl R (2007) Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev* 21(23):3027–3043.
11. Wachsmuth M, Caudron-Herger M, Rippe K (2008) Genome organization: Balancing stability and plasticity. *Biochim Biophys Acta* 1783(11):2061–2079.
12. Biggin MD, McGinnis W (1997) Regulation of segmentation and segmental identity by Drosophila homeoproteins: The role of DNA binding in functional activity and specificity. *Development* 124(22):4425–4433.
13. Hayashi S, Scott MP (1990) What determines the specificity of action of Drosophila homeodomain proteins?. *Cell* 63(5):883–894.
14. Lewis M, et al. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271(5253):1247–1254.
15. Yu J, Xiao J, Ren X, Lao K, Xie XS (2006) Probing gene expression in live cells, one protein molecule at a time. *Science* 311(5767):1600–1603.
16. Cornish PV, Ha T (2007) A survey of single-molecule techniques in chemical biology. *ACS Chem Biol* 2(1):53–61.
17. Xie XS, Choi PJ, Li GW, Lee NK, Lia G (2008) Single-molecule approach to molecular biology in living bacterial cells. *Ann Rev Biophys* 37:417–444.
18. Hager GL, McNally JG, Misteli T (2009) Transcription dynamics. *Mol Cell* 35(6):741–753.
19. Yao J, Ardehali MB, Fecko CJ, Webb WW, Lis JT (2007) Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. *Mol Cell* 28(6):978–990.
20. Yao J, Munson KM, Webb WW, Lis JT (2006) Dynamics of heat shock factor association with native gene loci in living cells. *Nature* 442(7106):1050–1053.
21. Yao J, Zobeck KL, Lis JT, Webb WW (2008) Imaging transcription dynamics at endogenous genes in living Drosophila tissues. *Methods* 45(3):233–241.
22. Papadopoulos DK, et al. Function and specificity of synthetic Hox transcription factors in vivo. *Proc Natl Acad Sci USA* doi:ADD DOI HERE.
23. Vukojević V, et al. (2008) Quantitative single-molecule imaging by confocal laser scanning microscopy. *Proc Natl Acad Sci USA* 105(47):18176–18181.
24. Ehrenberg M, Rigler R (1974) Rotational Brownian-motion and fluorescence intensity fluctuations. *Chemical Phys* 4(3):390–401.
25. Elson EL, Magde D (1974) Fluorescence correlation spectroscopy 1. Conceptual basis and theory. *Biopolymers* 13(1):1–27.
26. Magde D, Webb WW, Elson E (1972) Thermodynamic fluctuations in a reacting system—Measurement by fluorescence correlation spectroscopy. *Phys Rev Lett* 29(11):705–708.
27. Bacia K, Schwille P (2007) Fluorescence correlation spectroscopy. *Method Mol Cell Biol* 398:73–84.
28. Vukojević V, et al. (2005) Study of molecular events in cells by fluorescence correlation spectroscopy. *Cell Mol Life Sci* 62(5):535–550.
29. Elson EL (2001) Fluorescence correlation spectroscopy measures molecular transport in cells. *Traffic* 2(11):789–796.
30. Panzer S, Weigel D, Beckendorf SK (1992) Organogenesis in Drosophila melanogaster: embryonic salivary gland determination is controlled by homeotic and dorsoventral patterning genes. *Development* 114(1):49–57.
31. Berry M, Gehring W (2000) Phosphorylation status of the SCR homeodomain determines its functional activity: Essential role for protein phosphatase 2A,B′. *EMBO J* 19(12):2946–2957.
32. Gehring WJ, et al. (1994) Homeodomain–DNA recognition. *Cell* 78(2):211–223.
33. Sorensen JG, Nielsen MM, Kruhoffer M, Justesen J, Loeschcke V (2005) Full genome gene expression analysis of the heat stress response in Drosophila melanogaster. *Cell Stress Chaperon* 10(4):312–328.
34. Schneider S, Wunsch S, Schwab A, Oberleithner H (1996) Rapid activation of calcium-sensitive Na + /H+ exchange induced by 20-hydroxyecdysone in salivary gland cells of Drosophila melanogaster. *Mol Cell Endocrinol* 116(1):73–79.
35. Halford SE (2009) An end to 40 years of mistakes in DNA-protein association kinetics?. *Biochem Soc Trans* 37(Pt 2):343–348.
36. Mendes P (1997) Biochemistry by numbers: Simulation of biochemical pathways with Gepasi 3. *Trends Biochem Sci* 22(9):361–363.
37. Vukojević V, Sørensen PG, Hynne F (1996) Predictive value of a model of the Briggs–Rauscher reaction fitted to quenching experiments. *J Phys Chem-US* 100(43): 17175–17185.

Supporting Information

Vukojević et al. 10.1073/pnas.0914612107

SI Text

Supplement 1: Determination of DNA – Scr-HD Binding Constants in the Polytene Nucleus. ***Nonspecific interactions.*** Nonspecific interactions between DNA and Scr-HD$_{DD}$ can be described as a reversible reaction:

$$\mathrm{DNA} + \mathrm{ScrHD_{DD}} \underset{k_{-ns}}{\overset{k_{ns}}{\Leftrightarrow}} \mathrm{DNA - ScrHD_{DD}}, \qquad \textbf{[S1]}$$

where k_{ns} and k_{-ns} are rate constants for the formation and dissociation of the nonspecific complex between Scr-HD$_{DD}$ and DNA, respectively. At equilibrium the forward and backward reactions are equal:

$$k_{ns} \cdot [\mathrm{DNA}] \cdot [\mathrm{ScrHD_{DD}}] = k_{-ns} \cdot [\mathrm{DNA - ScrHD_{DD}}] \qquad \textbf{[S2]}$$

and the concentration of nonspecifically bound Scr-HD$_{DD}$ is:

$$[\mathrm{DNA - ScrHD_{DD}}] = \frac{k_{ns}}{k_{-ns}} \cdot [\mathrm{DNA}] \cdot [\mathrm{ScrHD_{DD}}]. \qquad \textbf{[S3]}$$

The total concentration of the bound and unbound reactive species is constant and equal to the respective initial concentrations. Therefore, concentrations of the reactive species at any time are:

$$[\mathrm{ScrHD_{DD}}] = [\mathrm{ScrHD_{DD}}]_0 - [\mathrm{DNA - ScrHD_{DD}}] \qquad \textbf{[S4]}$$

$$[\mathrm{DNA}] = [\mathrm{DNA}]_0 - [\mathrm{DNA - ScrHD_{DD}}] \qquad \textbf{[S5]}$$

The total concentration of potential binding sites in a polytene nuclei is noticeably higher than the total concentration of the transcription factor under physiologically relevant conditions. Changes in the DNA concentration due to Scr-HD$_{DD}$ binding can be therefore neglected and the DNA concentration can be regarded as constant and equal to the total DNA concentration:

$$[\mathrm{DNA}] \approx [\mathrm{DNA}]_0. \qquad \textbf{[S6]}$$

Introducing relations **S4** and **S5** in Eq. **S6** gives:

$$[\mathrm{DNA - ScrHD_{DD}}] = \frac{k_{ns}}{k_{-ns}} \cdot [\mathrm{DNA}]_0 \cdot ([\mathrm{ScrHD_{DD}}]_0 - [\mathrm{DNA - ScrHD_{DD}}]) \qquad \textbf{[S7]}$$

$$\left(1 + \frac{k_{ns}}{k_{-ns}} \cdot [\mathrm{DNA}]_0\right)[\mathrm{DNA - ScrHD_{DD}}] = \frac{k_{ns}}{k_{-ns}} \cdot [\mathrm{DNA}]_0 \cdot [\mathrm{ScrHD_{DD}}]_0 \qquad \textbf{[S8]}$$

$$[\mathrm{DNA - ScrHD_{DD}}] = \frac{\frac{k_{ns}}{k_{-ns}}}{\left(1 + \frac{k_{ns}}{k_{-ns}} \cdot [\mathrm{DNA}]_0\right)} \cdot [\mathrm{DNA}]_0 \cdot [\mathrm{ScrHD_{DD}}]_0 \qquad \textbf{[S9]}$$

$$[\mathrm{DNA - ScrHD_{DD}}] = \frac{k_{ns}}{(k_{-ns} + k_{ns} \cdot [\mathrm{DNA}]_0)} \cdot [\mathrm{DNA}]_0 \cdot [\mathrm{ScrHD_{DD}}]_0. \qquad \textbf{[S10]}$$

As expected, concentration of the nonspecific complex [DNA-Scr-HD$_{DD}$] increases linearly with the concentration of the transcription factor.

Using the data from FCS measurements (Fig. 5, *Black Line*), we estimate the dissociation constant for the nonspecific DNA – Scr-HD DD complex to be:

$$\frac{k_{ns} \cdot [\mathrm{DNA}]_0}{(k_{-ns} + k_{ns} \cdot [\mathrm{DNA}]_0)} = (0.35 \pm 0.15) \qquad \textbf{[S11]}$$

$$k_{ns} \cdot [\mathrm{DNA}]_0 = 0.35 \cdot (k_{-ns} + k_{ns} \cdot [\mathrm{DNA}]_0) \qquad \textbf{[S12]}$$

$$(1 - 0.3) k_{ns} \cdot [\mathrm{DNA}]_0 = 0.35 \cdot k_{-ns} \qquad \textbf{[S13]}$$

$$k_{ns} \cdot [\mathrm{DNA}]_0 = \frac{0.35}{(1 - 0.35)} \cdot k_{-ns} \qquad \textbf{[S14]}$$

$$k_{ns} \cdot [\mathrm{DNA}]_0 = 0.54 \cdot k_{-ns} \qquad \textbf{[S15]}$$

$$\frac{k_{-ns}}{k_{ns}} = K_d^{\mathrm{ScrHD_{DD},ns}} = \frac{[\mathrm{DNA}]_0}{0.54} \qquad \textbf{[S16]}$$

Nonspecific and Specific Interactions. The transcriptionally active variants of Scr-HD$_{wt}$ and Scr-HD$_{AA}$, undergo both specific and nonspecific interactions with DNA. Assuming that nonspecific interactions precede the specific ones, a two-step process of consecutive reactions was anticipated:

$$\mathrm{DNA_{ns}} + \mathrm{ScrHD_{wt}} \underset{k_{-ns}}{\overset{k_{ns}}{\Leftrightarrow}} (\mathrm{DNA - ScrHD_{wt}})_{ns}$$

$$(\mathrm{DNA - ScrHD_{wt}})_{ns} + \mathrm{DNA}_s \underset{k_{-s}}{\overset{k_s}{\Leftrightarrow}} (\mathrm{DNA - ScrHD_{wt}})_s \qquad \textbf{[S17]}$$

The turnover rate for the nonspecific complex is:

$$\frac{d[(\mathrm{DNA - ScrHD_{wt}})_{ns}]}{dt} = k_{ns} \cdot [\mathrm{DNA_{ns}}] \cdot [\mathrm{ScrHD_{wt}}] - (k_{-ns} + k_s \cdot [\mathrm{DNA}_s]) \cdot [(\mathrm{DNA - ScrHD_{wt}})_{ns}] + k_{-s} \cdot [(\mathrm{DNA - ScrHD_{wt}})_s] \qquad \textbf{[S18]}$$

Assuming a quasi-steady state approximation:

$$\frac{d[(\mathrm{DNA - ScrHD_{wt}})_{ns}]}{dt} = 0 \qquad \textbf{[S19]}$$

$$(k_{-ns} + k_s \cdot [\mathrm{DNA}_s]) \cdot [(\mathrm{DNA - ScrHD_{wt}})_{ns}] = k_{ns} \cdot [\mathrm{DNA_{ns}}] \cdot [\mathrm{ScrHD_{wt}}] + k_{-s} \cdot [(\mathrm{DNA - ScrHD_{wt}})_s]. \qquad \textbf{[S20]}$$

Using the mass balance equation to express the concentration of the free transcription factor:

$$[\mathrm{ScrHD_{wt}}] = [\mathrm{ScrHD_{wt}}]_0 - [(\mathrm{DNA - ScrHD_{wt}})_{ns}] - [(\mathrm{DNA - ScrHD_{wt}})_s] \qquad \textbf{[S21]}$$

and assuming as previously that:

$$[\mathrm{DNA}]_{ns} \approx [\mathrm{DNA}]_0, \qquad \textbf{[S22]}$$

Eq. **S23** becomes:

$$\begin{aligned}&(k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s])\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_{\mathrm{ns}}]\\&=k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0\cdot([\mathrm{ScrHD}_{\mathrm{wt}}]_0-[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_{\mathrm{ns}}]\\&-[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s])+k_{-s}\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s]\end{aligned}\qquad\textbf{[S23]}$$

$$\begin{aligned}&(k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s]+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0)\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_{\mathrm{ns}}]\\&=k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0\cdot([\mathrm{ScrHD}_{\mathrm{wt}}]_0-[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s])\\&+k_{-s}\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s]\end{aligned}\qquad\textbf{[S24]}$$

$$\begin{aligned}&(k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s]+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0)\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_{\mathrm{ns}}]\\&=k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0\cdot[\mathrm{ScrHD}_{\mathrm{wt}}]_0\\&-(k_{\mathrm{ns}}\cdot[\mathrm{DNA}]-k_{-s})\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s]\end{aligned}\qquad\textbf{[S25]}$$

$$\begin{aligned}[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_{\mathrm{ns}}]=&\frac{k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}{k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s]+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}\\&\cdot[\mathrm{ScrHD}_{\mathrm{wt}}]_0\\&-\frac{k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0-k_{-s}}{k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s]+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}\\&\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s].\end{aligned}\qquad\textbf{[S26]}$$

According to Eq. **S26** and the FCS data presented in Fig. 5 (*Red Stars*), the slope of the linear dependence gives:

$$\frac{k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}{k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s]+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}=(0.60\pm0.05)\qquad\textbf{[S27]}$$

and the intercept:

$$\begin{aligned}&\frac{k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0-k_{-s}}{k_{-\mathrm{ns}}+k_s\cdot[\mathrm{DNA}_s]+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s]\\&=(50\pm30)\ \mathrm{nM}.\end{aligned}\qquad\textbf{[S28]}$$

If k_{-s} is small compared to $k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0$ and can therefore be neglected, then:

$$0.6\cdot[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s]=50\ \mathrm{nM}.\qquad\textbf{[S29]}$$

Thus, the concentration of specific complex between *Scr*-HD wild type and DNA in the polytene nucleus can be estimated to be:

$$[(\mathrm{DNA}-\mathrm{ScrHD}_{\mathrm{wt}})_s]=(80\pm50)\ \mathrm{nM}.\qquad\textbf{[S30]}$$

FCS data enabled us also to estimate the dissociation constant for the nonspecific DNA – Scr-HD wt complex (Fig. 5, *Dotted Red Line*):

$$\frac{k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0}{(k_{-\mathrm{ns}}+k_{\mathrm{ns}}\cdot[\mathrm{DNA}]_0)}=(0.65\pm0.05)\qquad\textbf{[S31]}$$

$$\frac{k_{-\mathrm{ns}}}{k_{\mathrm{ns}}}=K_d^{\mathrm{ScrHD}_{\mathrm{wt,ns}}}=\frac{[\mathrm{DNA}]_0}{1.9}.\qquad\textbf{[S32]}$$

Supplement 2: Calculation of the Concentration of DNA Sites for Specific and Nonspecific Binding of the Scr-HD Within a Polytene Nucleus. The Drosophila haploid genome size is 1.76×10^8 bp long; this means that the diploid genome size is $2\times1.76\times10^8=3.52\times10^8$ bp long.

Polytene cells normally undergo 10 DNA-replication cycles without cell division, thus $2^{10}=1024$ chromosomal copies are produced.

This results in $1024\times3.52\times10^8=3.6\times10^{11}$ bp within a polytene nucleus.

However, a protein molecule can interact (bind specifically or nonspecifically) only with a few base-pairs of the DNA.

A) For nonspecific binding (i.e., binding to DNA sequences that do not belong to the putative *cis*-regulatory elements bound by Scr) we assume on the average a 6 bp DNA interaction domain with the consensus sequence TAATCG or TAATGG found for Scr, Unx, and Antp (1–4). Taking into consideration that this consensus occurs every 2048 bp in the genome, this results in $\frac{3.6\times10^{11}}{2048}=1.758\times10^8$ sites within a polytene nucleus. 1 mol of potential sites contains 6.023×10^{23} sites.

n mol of potential sites contain 1.758×10^8 sites, thus $n=\frac{1.758\times10^8}{6.023\times10^{23}}=0.292\times10^{-15}$ mol of nonspecific sites within the polytene nucleus.

However, the Drosophila euchromatin comprises 117 Mbp (5), which means that only 66.5% of the total DNA length can be considered as candidate sequences for transcription factor–DNA interactions. This results in $0.292\times10^{-15}\times0.665=0.194\times10^{-15}$ mol of binding sites.

Experimentally we have determined the diameter of a polytene nucleus as 20 μm, thus the radius is 10 μm. The volume of the nucleus will be

$$\begin{aligned}V_{\mathrm{nucleus}}=\frac{4}{3}\pi r^3=\frac{4}{3}\times3.14\times(10\times10^{-6})^3\ \mathrm{m}^3=4.187\times10^{-15}\ \mathrm{m}^3\\=4.187\times10^{-15}\times10^3\ (\mathrm{dm})^3=4.187\times10^{-12}\ \mathrm{L}.\end{aligned}\qquad\textbf{[S33]}$$

The concentration of the potential DNA-transcription factor binding sites will be thus

$$C_{\mathrm{sites}}^{\mathrm{non-specific}}=\frac{n}{V_{\mathrm{nucleus}}}=\frac{1.94\times10^{-16}}{4.187\times10^{-12}}=46\ \mu\mathrm{M}.\qquad\textbf{[S34]}$$

B) For specific binding to occur, usually a larger sequence is necessary. One example are the *fkh250* and *fkh250*$^{\mathrm{con}}$ elements, which consist of 10 bp (AGATTAATCG) and can bind either an Scr-Exd or a Ubx-Exd complex, respectively (6). Assuming thus a minimum requirement of 10 bp to confer Scr binding specificity, this sequence may occur once every $4^{10}=1048576\cong1.048\times10^6$ bp in the genome.

This results in $\frac{3.65\times10^{11}}{1.048\times10^6}=3.48\times10^5$ sites within a polytene nucleus.

1 mol of potential sites contains 6.023×10^{23} sites.

n mol of potential sites contain 3.48×10^5 sites thus $n=\frac{3.48\times10^5}{6.023\times10^{23}}=0.578\times10^{-18}$ mol.

Thus the concentration of specific binding sites within the polytene nucleus will be

$$C_{\mathrm{sites}}^{\mathrm{specific}}=\frac{n}{V_{\mathrm{nucleus}}}=\frac{0.578\times10^{-18}\times0.665}{4.187\times10^{-12}}=92\ \mathrm{nM}.\qquad\textbf{[S35]}$$

Supplement 3: Isoelectric Points of *Scr*-HD Variants. Isoelectric points for *Scr*-HD variants fused with mCitrine were estimated using the Protein Calculator v3.3 that is available on-line (http://www.scripps.edu/cgi-bin/cdputnam/protcalc3). The primary sequences given below were submitted and the charge at different pH was estimated by the program (Table S1).

mCitrine-*Scr*-HD$_{\mathrm{wt}}$

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG-DATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLMCFARY-PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAE-VKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHN-

VYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNT-
PIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVT-
AAGITLGMDELYKLEPPQIYPWMKRVHLGTSTVNANGE-
TKRQRTSYTRYQTLELEKEFHFNRYLTRRRRIEIAHALC-
LTERQIKIWFQNRRMKWKKEHKMASMNIVPYHMGPYG-
HPYHQFDIHPSQFAHLSA

mCitrine-Scr-HD$_{AA}$

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG-
DATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLMCFARY-
PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAE-
VKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNV-
YIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPI-
GDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTA-
AGITLGMDELYKLEPPQIYPWMKRVHLGTSTVNANGET-
KRQRAAYTRYQTLELEKEFHFNRYLTRRRRIEIAHALCL-
TERQIKIWFQNRRMKWKKEHKMASMNIVPYHMGPYG-
HPYHQFDIHPSQFAHLSA

mCitrine-Scr-HD$_{DD}$

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG-
DATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLMCFAR-
YPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRA-
EVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSH-
NVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQN-
TPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEF-
VTAAGITLGMDELYKLEPPQIYPWMKRVHLGTSTVNA-
NGETKRQRDDYTRYQTLELEKEFHFNRYLTRRRRIEI-
AHALCLTERQIKIWFQNRRMKWKKEHKMASMNIVPY-
HMGPYGHPYHQFDIHPSQFAHLSA

mCitrine-Scr-HD$_{DD}^{Q50AN51A}$

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEG-
DATYGKLTLKFICTTGKLPVPWPTLVTTFGYGLMCFAR-
YPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTR-
AEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNS-
HNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQ-
NTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLE-
FVTAAGITLGMDELYKLEPPQIYPWMKRVHLGTSTVNA-
NGETKRQRDDYTRYQTLELEKEFHFNRYLTRRRRIEIA-
HALCLTERQIKIWFAARRMKWKKEHKMASMNIVPYHM-
GPYGHPYHQFDIHPSQFAHLSA

Supplement 4: Materials and Methods. ***Immunohistochemistry.*** Salivary glands from third instar wandering larvae were dissected at room temperature (RT) by pulling the mouth-hooks from the larval body in Grace's medium (Gibco) and and the whole tissue was fixed for 10 min in Grace's medium containing 5.3% EM-grade paraformaldehyde (Electron Microscopy Science, PA) freshly used every 10 days. After twice washed for 15 min in PBT (PBS containing 0.1% TritonX-100) and blocked for at least 30 min in PBT containing 5% normal goat serum (NGS) and 20 μg/mL NaN_3, samples were incubated with primary antibodies in blocking solution overnight. Subsequently, samples were washed 3 times for 20 min in PBT and secondary antibodies were applied in PBS for 4 h at RT. After 3 washes for 20 min with PBT and equilibration in PBS, the salivary glands were separated from the mouth-hooks and mounted in Vectashield (Vector Laboratories). Images were obtained within 24 h using a Leica SP5 Confocal setup.

Purification of mCitrine-Scr-HD$_{wt}$. mCitrine-Scr-HD$_{wt}$ was cloned as an *Eco*RI-*Not*I fragment into the pET-21b(+) vector (Novagen), and the 6× histidine-tagged peptide produced in Rosetta™2 (DE3)pLysS cells (Novagen). Purification of a small amount of recombinant protein was performed using the ProBond™ Purification System (Invitrogen). Expression of the peptide was confirmed by SDS-PAGE and Western blotting using a mouse anti-GFP antibody (Molecular Probes). Protein was concentrated and used for FCS analysis in solution.

Preparation of salivary glands for FCS and heat shock procedures. For FCS measurements 3–4 pairs of salivary glands were dissected from third instar wandering larvae in Grace's medium and transferred to chambered slides (Nalge Nunc International) containing PBS. For titrating the concentration of Scr-HD in the nucleus, variable (from 10 min up to 2 h) heat-shocks were applied to the larvae prior to dissection. Then, larvae were dissected within short time and used in the measurements.

APD imaging. An individually modified instrument (Zeiss, LSM 510, ConfoCor 3) with fully integrated FCS/CLSM optical pathways was used for imaging. The detection efficiency of CLSM imaging was significantly improved by the introduction of APD detectors. As compared to PMTs, which are normally used as detectors in conventional CLSM, the APDs are characterized by higher quantum yield and collection efficiency—about 70% in APDs as compared to 15–25% in PMTs, higher gain, negligible dark current and better efficiency in the red part of the spectrum. Enhanced fluorescence detection efficiency enabled image collection using fast scanning (1–5 μs/pixel). This enhances further the signal-to-noise-ratio by avoiding fluorescence loss due to triplet state formation, enabling fluorescence imaging with single-molecule sensitivity. In addition, low laser intensities (150–750 μW) could be applied for imaging, significantly reducing the photo-toxicity. Advantages of APD imaging for nondestructive observation of molecular interactions in real time in live cells are described in detail in reference 23, listed in the main text.

Short background on fluorescence correlation spectroscopy (FCS). FCS measurements are performed by recording fluorescence intensity fluctuations in a very small, approximately ellipsoidal volume element (about 400 nm wide and 1.5 μm long) that is generated in the salivary gland nuclei by focusing the laser light through the objective of the microscope. The fluorescence intensity fluctuations, caused by fluorescently labeled molecules passing through the volume element, are analyzed using statistical methods. Temporal autocorrelation analysis was used in this study, but other methods can be also applied. For recent reviews on FCS see references 27–29 in the main text.

In temporal autocorrelation analysis we first derive the autocorrelation function $G(t)$:

$$G(\tau) = 1 + \frac{\langle \delta I(t) \delta I(t+\tau) \rangle}{\langle I \rangle^2}, \qquad \textbf{[S36]}$$

that gives the correlation between the intensity of light, $I(t)$, measured at a certain time, t, and its intensity, $I(t+\tau)$, measured at a later time $t+\tau$. For further analysis, an autocorrelation curve is derived by plotting $G(\tau)$ as a function of different lags, i.e. different autocorrelation times τ. See for example Figs 2–4 in the main text. The amplitude of the autocorrelation function is reciprocally proportional to the average number of molecules in the observation volume. Thus, the amplitude of the autocorrelation curve decreases for increasing number of molecules (Fig. 2*D*, Fig. 4*A* and *B*). The inflection points on the autocorrelation curves reflect the characteristic times for the investigated molecular processes.

The experimentally obtained autocorrelation curves are compared to autocorrelation functions derived for different model systems. A model describing free diffusion of two components and triplet formation was used in this study:

$$G(\tau) = 1 + \frac{1}{N} \cdot \left(\frac{1-y}{(1+\frac{\tau}{\tau_{D1}})\sqrt{1+\frac{w_{xy}^2}{w_z^2}\frac{\tau}{\tau_{D1}}}} + \frac{y}{(1+\frac{\tau}{\tau_{D2}})\sqrt{1+\frac{w_{xy}^2}{w_z^2}\frac{\tau}{\tau_{D2}}}} \right) \cdot \left[1 + \frac{T}{1-T} \exp\left(-\frac{\tau}{\tau_T} \right) \right] \qquad \textbf{[S37]}$$

In the above equation, N is the average number of molecules in the observation volume element, y is the fraction of the slowly moving Scr-HD molecules, τ_{D1} is the diffusion time of the free Scr-HD molecules, τ_{D2} is the diffusion time of Scr-HD molecules undergoing interactions with the DNA, w_{xy} and w_z are radial and axial parameters, respectively, related to spatial properties of the detection volume element, T is the average equilibrium fraction of molecules in triplet state and τ_T the triplet correlation time, related to rate constants for intersystem crossing and the triplet decay. Spatial properties of the detection volume, represented by the square of the ratio of the radial and axial parameters ($(w_{xy}/w_z)^2$), are determined in calibration measurements performed by using a solution of a fluorescent dye Rhodamine 6G for which the diffusion coefficient (D) is known. The diffusion time, τ_D, measured by FCS, is related to the translation diffusion coefficient D by:

$$\tau_D = \frac{w_{xy}^2}{4D}. \qquad \text{[S38]}$$

1. Ekker SC, et al. (1994) The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. *EMBO J* 13(15):3551–3560.
2. Ekker SC, Young KE, von Kessler DP, & Beachy PA (1991) Optimal DNA sequence recognition by the Ultrabithorax homeodomain of Drosophila. *EMBO J* 10(5):1179–1186.
3. Muller M, et al. (1988) Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J* 7(13):4299–4304.
4. Ryoo HD & Mann RS (1999) The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* 13(13):1704–1716.
5. Hoskins RA, et al. (2002) Heterochromatic sequences in a Drosophila whole-genome shotgun assembly. *Genome Biol* 3(12):RESEARCH0085.
6. Joshi R, et al. (2007) Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* 131(3):530–543.
7. Schneider S, Wunsch S, Schwab A, & Oberleithner H (1996) Rapid activation of calcium-sensitive Na + /H+ exchange induced by 20-hydroxyecdysone in salivary gland cells of Drosophila melanogaster. *Mol Cell Endocrinol* 116(1):73–79.

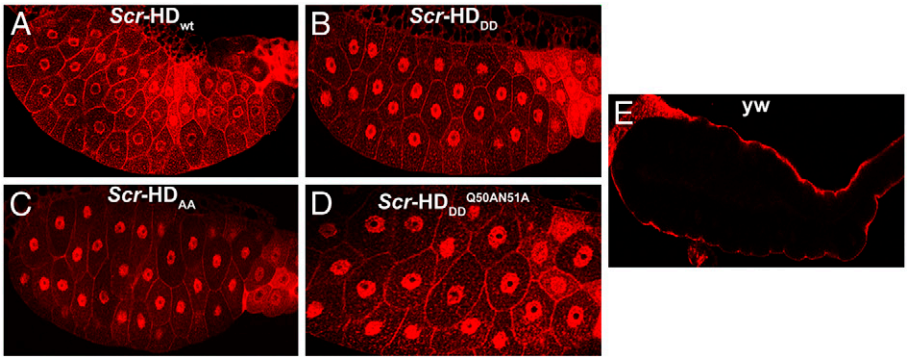

Fig. S1. Flies carrying UAS constructs but no Gal4 driver constructs express the transgenes in small amounts at normal conditions. (*A–E*) Immunostaining of salivary glands of third instar wandering larvae raised at 25 °C, carrying UAS–mCitrine–*Scr*-HD constructs (without any Gal4 driver). All glands exhibit nuclear localization of the synthetic transcription factor (*A–D*), as compared to the *yw*^{1118} control (injection background) (*E*). The transcription factor seemingly follows patterned nuclear distribution among all variants, due to paraformaldehyde fixation and cross-linking with the DNA. The tight association with nuclear DNA persists for Scr-HD$_{wt}$ and Scr-HD$_{AA}$, but not for Scr-HD$_{DD}$ or Scr-HD$_{DD}^{Q50AN51A}$, when visualized in live cells by APD imaging.

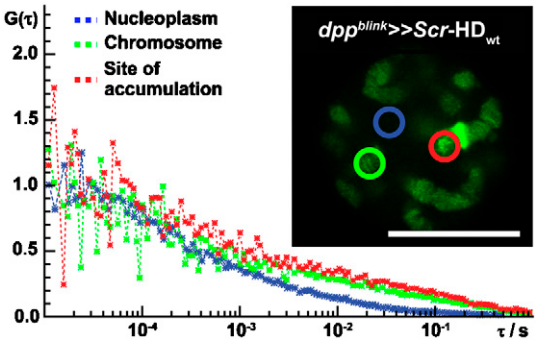

Fig. S2. DNA – Scr-HD$_{wt}$ interactions in different regions of the polytene nucleus. APD imaging of nuclei overexpressing *Scr*-HD$_{wt}$ showed differential nuclear distribution of the transcription factor in the nucleoplasm (*Blue Circle*); on the chromosomes (*Green Circle*) and in sites of accumulation (*Red Circle*). The dynamics of Scr-HD$_{wt}$ is different in the corresponding regions, reflecting differences in its interactions with DNA. Scale bar is 20 μm.

Table S1. Isoelectric points for Scr-HD variants at different pH

	Charge			
pH	mCitrine-Scr-HD_{wt}	mCitrine-Scr-HD_{AA}	mCitrine-Scr-HD_{DD}	mCitrine-Scr-$HD_{DD}^{Q50AN51A}$
7.00	7.2	7.2	5.2	5.2
7.10	6.4	6.4	4.4	4.4
7.20*	**5.7**	**5.7**	**3.7**	**3.7**
7.30*	**5.1**	**5.1**	**3.1**	**3.1**
7.40*	**4.5**	**4.5**	**2.5**	**2.5**
7.50*	**4.0**	**4.0**	**2.0**	**2.0**
7.60*	**3.5**	**3.5**	**1.5**	**1.5**
7.70*	**3.1**	**3.1**	**1.1**	**1.1**
7.80*	**2.7**	**2.7**	**0.7**	**0.7**
7.90*	**2.3**	**2.3**	**0.3**	**0.3**
8.00	1.9	1.9	-0.1	-0.1
8.10	1.5	1.5	-0.5	-0.5
8.20	1.0	1.0	-1.0	-1.0
8.30	0.5	0.5	-1.5	-1.5
8.40	0.0	0.0	-2.0	-2.0
8.50	-0.5	-0.5	-2.5	-2.5
8.60	-1.1	-1.1	-3.1	-3.1
8.70	-1.8	-1.8	-3.8	-3.8
8.80	-2.6	-2.6	-4.6	-4.6
8.90	-3.4	-3.4	-5.4	-5.4

*Indicated values correspond to experimentally measured pH in salivary gland cells isolated from third instar larvae (7).

FUNCTIONAL SYNTHETIC *ANTENNAPEDIA* GENES: DUAL ROLE OF THE YPWM MOTIF IN TRANSCRIPTIONAL ACTIVATION AND REPRESSION

Dimitrios K. Papadopoulos[1,*], Diana Reséndez-Pérez[2,*], Diana L. Cárdenas-Chávez[2], Karina Villanueva-Segura[2], Ricardo Canales del Castillo[2], Daniel A. Felix[1,3], Raphael Fünfschilling[1], and Walter J. Gehring[1,#]

[1]Department of Cell Biology, Biozentrum, University of Basel, 4056 Basel, Switzerland

[2]Facultad de Ciencias Biológicas UANL, Ciudad Universitaria, C.P. 66450, Monterrey, N.L., México

[3]Present address: Institute of Genetics, Queen's Medical Centre, University of Nottingham, Nottingham, UK

[*]These authors contributed equally to this work

[#]Correspondence: walter.gehring@unibas.ch

Abstract

Segmental identity along the anteroposterior axis of bilateral animals is governed by *Hox* genes, encoding transcription factors that harbor the homeodomain, a DNA-binding domain conserved throughout evolution and among different *Hox* paralogs. Although structural and genetic studies have identified cofactors interacting with the conserved Hox YPWM motif, our understanding about Hox specificity remains incomplete. We have previously shown that synthetic *Sex combs reduced* (*Scr*) genes are functional *in vivo* and the long N-terminal sequences of Scr not required for homeotic function in the fly. Here, we extend this knowledge to synthetic *Antennapedia* (*Antp*) genes and analyze their function throughout development. Antp peptides lacking the entire N-terminus of the protein caused head involution defects in the larval cuticle and tarsal transformations in adult antennae, repressed embryonic (*Scr*) and antennal (*Spalt major*, *distal antenna*) genes and ectopically activated *Teashirt* and *grain* in the respective tissues. Their transactivation capacity was altered in the presence of Extradenticle and Bip2 in a YPWM-dependent fashion, suggesting a function comparable to the full-length homeoprotein. Also, the Antp YPWM motif was found an essential requirement for transcriptional activation, while its substitution by the WRPW repressor motif still allowed the protein to act as a repressor. Finally, variations in linker size, naturally occurring in Antp isoforms, enhanced selectively its repressive or activating efficiency, suggesting a role of the linker in target selection. Taken together, our results exempt the N-terminal portion of Antp from a role in specificity, indicate the dual role of the YPWM motif in transcriptional regulation and propose synthetic genes as a potential tool for synthetic biology, developmental studies and biomedical applications.

Introduction

Homeotic (*Hox*) genes encode transcription factors that differentially specify body segments along the anterior-posterior axis of bilateral animals (Gehring, 1987; Lewis, 1978; McGinnis and Krumlauf, 1992). Hox transcription factors harbour the homeodomain (HD), a helix-turn-helix DNA-binding domain that has been highly conserved throughout animal evolution (Gehring et al., 2009; Kmita-Cunisse et al., 1998) and shares remarkable sequence similarity among different Hox paralogs (McGinnis et al., 1984c). In order for Hox proteins to confer differential identity in the developing embryo, they need to regulate differentially – notwithstanding their sequence similarity – a very high number of target genes, most of the times in a paralog-specific manner. Despite our current knowledge of their DNA binding properties (Gehring et al., 1994); the identification of Hox cofactors and collaborators that increase their target-selection precision – reviewed by (Mann et al., 2009); the possibility to perform genome-wide analyses – for reviews see for example (Hueber and Lohmann, 2008; Pearson et al., 2005); and our understanding of transcriptional regulation *in vivo*; we fall short of providing a complete explanation about the specificity of Hox factors, such that would lead to the prediction and/or identification of novel Hox functions.

The lack of methodologies that would provide once-and-for-all an answer to the fundamental question of specificity has given rise to individual studies which aimed at dissecting the molecular properties of Hox peptides, structurally or genetically. Recent milestones in *Hox* research have been for example: 1) the finding that TALE (three amino acid loop extension) HD proteins, like Extradenticle (Exd) and Homothorax (Hth) in the fly, bind the Hox YPWM motif *in vitro* (Johnson et al., 1995), and behave as putative Hox cofactors *in vivo* (Mann and Chan, 1996); 2) the molecular resolution of the DNA-bound Ultrabithorax (Ubx)-Exd (Passner et al., 1999; Rohs et al., 2009), Scr-Exd (Joshi et al., 2007) and HoxB1-Pbx1 (Piper et al., 1999) complexes; 3) the realization that *Hox* specificity can be owed, at least partially, to DNA structure, rather than sequence (Joshi et al., 2007; Rohs et al., 2009); 4) the finding that *Hox* function can be cofactor independent (Peifer and Wieschaus, 1990; Rawat et al., 2008b; Shen et al., 2004b); 5) the fact that the lack of the YPWM motif does not completely abolish Hox function *in vivo* (Galant et al.,

2002; Merabet et al., 2003; Shen et al., 1997b). These findings have contributed immensely to our current perception of *Hox*-mediated gene regulation.

We have previously demonstrated that DNA-binding and the functional specificity of Antp reside in the HD and its flexible N-terminal arm, respectively (Furukubo-Tokunaga et al., 1993; Gehring et al., 1994; Qian et al., 1994b), as well as that the YPWM motif provides a direct link between Antp and the transcriptional machinery (Prince et al., 2008). In the present study, we use synthetic *Antp* genes, which encode a small portion of the full-length protein (spanning the region between the YPWM motif and the beginning of the C-terminus), to examine their function *in vivo*, throughout development. We show that the HD and the YPWM motif suffice for homeotic function in the fly and that the N-terminus of the wild type protein is dispensable for all described *Antp*-specific phenotypes. We establish that this conserved tetrapeptide is an absolute requirement for transcriptional activation, whereas in the cases of Antp functioning as a repressor, the YPWM motif can be substituted by WRPW, a domain interacting with the generic transcriptional corepressor groucho (gro) (Paroush et al., 1994), without loss of *Antp*-specificity *in vivo*. In addition, the linker region between the YPWM motif and the HD – comprised of 8 or 4 amino acids in different naturally occurring Antp isoforms (Stroeher et al., 1988) – was found to play an important role in transcriptional potency and/or specificity *in vivo*. Finally, we compare the transactivation efficiency of full-length and synthetic Antp peptides on consensus HD binding sites (BS2) and monitor the changes of their transcriptional behaviour in the presence of cofactors that interact with the YPWM motif.

Results

Synthetic *Antp* genes cause head involution defects in the embryonic cuticle and act as repressors and activators of genes required for normal embryonic development. In order to gain insight into the capacity of synthetic *Antp* constructs to trigger gene-specific phenotypes, we first examined their homeotic function during embryonic development. Abnormalities in the development of anterior (head) structures of the embryo, due to the transformation of head segments towards second thoracic segment identity (T2), are caused upon ectopic expression

of *Antp* in the segments anterior to its normal expression domain during development (Furukubo-Tokunaga et al., 1993; Zeng et al., 1993). The acquisition of mesothoracic fate in the embryonic head is manifested through the repression of prothoracic (T1) beards, normally specified by *Scr*, and the formation of ectopic denticle bands, normally specified by *Antp*, in the head segments (Furukubo-Tokunaga et al., 1993; Zeng et al., 1993). Using the *nullo*-Gal4 driver (Gehring et al., 2009) for strong and long-lasting expression of the *Antp* genes throughout the embryo, we observed that all synthetic and full-length constructs featuring either a YPWM motif or a WRPW repressor motif (Fig. 1 A-B, E-I and L-M) caused head involution defects of variable strength and repression of the prothoracic beards. In contrast, full-length ^{AAAA}Antp, bearing a substitution of the YPWM motif by alanines, exhibited a wild type-like phenotype (Fig. 1 C) and no repression of T1 beards was visible (Fig. 1 N), whereas expression of the synthetic ^{AAAA}*Antp* constructs (Fig. 1 J-K) resulted in mild defects, which appeared marginally enhanced for the short linker gene (Fig. 1 K). No repression of T1 identity was visible in embryos overexpressing these constructs (Fig. 1 O-P). The most dramatic phenotype was displayed by the full-length Antp with C-terminal fusions of WRPW (Fig. 1 B and D) and remarkable differences were observed between the synthetic ^{YPWM}Antp long and short linker peptides (Fig, 1 H-I). Finally, the addition of *en* repressor domain to the C-terminus of the full-length Antp (Fig. 1 E) did not result in any phenotype considerably different from the one caused by the wild type Antp protein.

Moreover, target genes repressed or activated by *Antp* in the embryo were used as a system to determine the ability of synthetic and full-length Antp peptides to act, directly or indirectly, as transcriptional repressors and activators. Such genes include *Scr* (Andrew et al., 1994), normally expressed in the labial and first thoracic segment of the embryo, and *Teashirt* (*Tsh*), expressed in the trunk in parasegments 3-13 (McCormick et al., 1995). Using the *patched* (*ptc*) enhancer (Brand and Perrimon, 1993) for ectopic *Antp* expression in the anterior region of every parasegment, we observed repression of *Scr* by the full-length and synthetic *Antp* lines in the labial and T1 segments (Fig. 2 A), ranging from partial to complete. Interestingly, Antp full-length and synthetic peptides bearing a YPWM to AAAA substitution did not repress *Scr*. A clear difference in repression potency was again observed between the synthetic ^{YPWM}*Antp* long and short linker genes, with the short linker peptide exhibiting enhanced repressor activity over its long linker counterpart.

Meanwhile, synthetic ^{WRPW}Antp peptides conferred only partial repression. However, unlike the repression of *Scr*, where the presence of a repressor motif (WRPW) seemed to suffice for functionality of the transcription factor, full-length and synthetic Antp peptides lacking a functional YPWM motif failed to activate ectopic expression of *Tsh* in the region of the embryonic head, suggesting an indispensable requirement of the YPWM motif in transcriptional activation (Fig. 2 B). Differences in the potency of ectopic *Tsh* activation was again observed between the synthetic ^{YPWM}*Antp* genes, but the trend was opposite from the one displayed in homeotic transformations in the embryo, or in the repression of *Scr*. This time the long linker construct activated *Tsh* in a domain as extensive as the full-length wild type peptide, whereas the effect of the short linker construct was weaker.

Synthetic *Antp* genes cause tarsal transformations of the adult antenna, similar to the full-length protein. Though described as *Antp*-specific phenotypes, the embryonic head involution defects do not suffice to describe the function of *Antp* in transcriptional regulation, since it is not known whether they are caused by *Antp* acting as a repressor, activator, or both. Therefore we used the *Dll* enhancer to generate tarsal transformations in the antenna. Ectopic T2-tarsi, bearing apical and/or preapical bristles, were generated only by full-length and synthetic Antp peptides featuring a functional YPWM motif (Fig. 3 A-B, E, H-I and L-M). As anticipated, no transformation was observed with the full-length ^{AAAA}Antp (Fig. 3 C), but the corresponding synthetic peptides conferred a mild aristal thickening and resulted in ectopic leg bristles on the third antennal segment (A3) (Fig. 3 J-K), suggesting mild, yet substantial, transformation. The same rule applied for ^{AAAA}Antp^{WRPW}, where numerous tarsal bristles and absence of the arista were identified on the antenna (Fig. 3 D). Synthetic and full-length Antp peptides with YPWM-to-WRPW substitutions caused overgrowth of the A3 segment (Fig. 3 F-G and L-M) which occasionally displayed leg bristles (Fig. 3 L). Finally, the antennal tarsi generated by the synthetic ^{YPWM}*Antp* lines (Fig. 3 H-I) showed a strong dependence on the linker region between the HD and YPWM motif, with the long linker construct (Fig. 3 H) giving rise to tarsi almost as dramatic as the full-length wild type peptide (Fig. 3 A).

Synthetic *Antp* genes repress antennal-specific and activate leg-specific genes in antennal discs undergoing tarsal transformations. Tarsal transformations in the antennal disc require the repression of genes needed for

antennal morphogenesis and ectopic activation of leg-specific genes, which will allow antennal disc cells to commit themselves to the leg-determination program (Gehring et al., 2009). We and others have shown previously that antennal-specific genes such as *Spalt major* (*Salm*) (Wagner-Bernholz et al., 1991) and *distal antenna* (*dan*) (Emerald et al., 2003) are repressed in the antennal disc, while leg-specific markers such as *grain* (*grn*) (Grieder et al., 2009; Wagner-Bernholz et al., 1991) are ectopically activated upon expression of *Antp*. The fact that full-length or synthetic *Scr* genes (Papadopoulos et al., 2010) also confer the aforementioned changes in the antennal determination program suggests a general mechanism, by which the regulation of *Salm*, *dan* and *grn* is altered in discs undergoing tarsal transformations. As expected, *Salm* (Fig. 4 A-Z) and *dan* (Fig. 4 A'-Z') were completely repressed in all *Antp* lines tested, with the exception of lines bearing the AAAA substitution of the YPWM motif (Fig. 4 E-F, S-T and U-V; E'-F', S'-T' and U'-V') and in some cases the ^{WRPW,LL}Antp and the ^{WRPW,SL}Antp synthetic peptides (Fig. W-Z and W'-Z'). These antennal discs were largely overgrown as compared to the eye portions of the disc and gave rise to an enlarged third antennal (A3) segment in adult flies (Fig. 3 L-M), but no pronounced tarsal phenotypes were observed. On the other hand, the ectopic activation of *grn* in the antennal disc (Fig. 5) followed the same principle as the ectopic activation of *Tsh* in the embryo: an absolute requirement for a functional YPWM motif (Fig. 5 B-C, F and I-J) was derived from comparison between all lines used, while the rest of the constructs were indistinguishable from the control disc, which bore a single enhancer trap chromosome (Fig, 5 A). Table S1 summarizes in a quantitative manner the phenotypes observed.

Synthetic Antp peptides activate transcription via YPWM-dependent Exd and Bip2 interactions. To describe quantitatively the ability of full-length and synthetic Antp peptides to bind DNA and activate transcription in live cells, we analyzed the transactivation efficiency of all constructs, using a luciferase reporter gene under the control of consensus HD binding sites (BS2) (Fig. 6 A). Co-transfection of the latter with wild type *Antp* showed strong activation of the reporter, whereas ^{AAAA}Antp increased luciferase activity about 50% (Fig. 6 B). In contrast, the WRPW motif behaved as a potent repressor domain, decreasing transactivation efficiency at least 50%, as compared to the wild type protein (Fig. 6 B). We observed this trend in both the C-terminal fusion of WRPW and the substitutions of the YPWM motif with WRPW. This result is consistent with previous findings using a Gal4-

WRPW protein, which decreased transactivation to almost 65% of the wild type protein (Fisher et al., 1996). Interestingly, the C-terminal fusion of WRPW to ^{AAAA}Antp behaved in a similar way as the ^{AAAA}Antp (Fig. 6B). *Antp* synthetic constructs also activated the reporter gene transcription with an efficiency of 30 to 50% of the full-length peptide (Fig. 6 B), while in all cases the long linker genes displayed a higher reporter induction than short linker genes (Fig. 6 C). As expected, synthetic *Antp* genes bearing a YPWM to alanines or WRPW substitution showed lower levels of reporter activity (50-80%) compared to the synthetic ^{YPWM,LL}Antp (Fig. 6 C).

Since synthetic Antp peptides might regulate transcription of target genes by interacting with putative Antp cofactors, such as Exd (Passner et al., 1999; Piper et al., 1999; Ryoo and Mann, 1999) and Bip2 (Gangloff et al., 2001; Prince et al., 2008), we carried out similar transactivation experiments including Exd (amino acids 144-376) or Bip2 (residues 2-89) (Fig. 7A). Our results showed that the transactivation efficiencies of synthetic ^{YPWM,LL}Antp and ^{YPWM,SL}Antp decreased approximately 20% – and increased almost 70% – respectively, in the presence of Exd (Fig. 7B). Contrary to these observations, the corresponding substituted synthetic peptides (Synth^{AAAA,LL}Antp and Synth^{AAAA,SL}Antp) did not show any change in transcriptional activity. In the same way, the presence of Bip2 decreased the transactivation potency of synthetic Antp-peptides up to 30%, while the synthetic AAAA peptides were not able to respond to Bip2 (Fig. 7C). These findings suggest that the synthetic Antp peptides may interact with these cofactors in a YPWM-dependent fashion.

Discussion

In the present study we have demonstrated that synthetic *Antp* genes are functional *in vivo*, since their homeotic function was found comparable to their full-length counterpart using various *in vivo* assays. We have previously shown that synthetic *Scr* genes encoding only a small C-terminal portion of *Scr*, containing the HD, exhibit homeotic function in the fly, which was found enhanced compared to the native fly protein (Papadopoulos et al., 2010), and used them as a tool to study their molecular interactions with chromosomal DNA in live cells (Vukojevic et al., 2010). Here we extended this notion to *Antp* and investigated the requirement of the YPWM

motif in transcriptional regulation *in vivo*. We have also deduced important conclusions about the function of the YPWM motif in Antp transcriptional regulation, mediated by full-length and synthetic peptides, and found a requirement for the tetrapeptide in both activation and repression. Substitutions of the YPWM motif by WRPW retained the repressive activity of Antp and C-terminal fusions of the WRPW motif to full-length peptides converted them into enhanced repressors. On the other hand, the *engrailed* (*en*) repression domain – also able to interact with Gro *in vivo* (Tolkunova et al., 1998) – did not confer significant changes to Antp functioning as a repressor or activator. We have also evaluated the role of the linker region between the YPWM motif and the HD, using the naturally occurring splicing variants of *Antp* with an 8 or a 4 amino-acid linker. We found that the long-linker variant is a stronger activator and the short-linker one a stronger repressor of transcription *in vivo*, a result which suggests that specificity might also depend on linker size. In transactivation experiments in *Drosophila* S2 cells we have indicated that the presence of the YPWM motif in full-length, but not synthetic, Antp peptides has a negative effect in BS2 element-mediated transcription, further supporting the notion that the N-terminus has a role in transcriptional potency, and not specificity, *in vivo*. Using the same system we have also confirmed our phenotypic findings that the WRPW motif behaves as a strong repressor motif in Antp peptides. Finally, we have monitored changes in transactivation levels upon coexpression of synthetic peptides with putative Antp cofactors, such as Exd and Bip2, in S2 cells. We have reported that in the presence of these cofactors the activity of synthetic Antp peptides, bearing a YPWM motif, changes significantly, a result suggesting YPWM-dependent interactions, as described previously for the full-length protein. Our results, while proving the functionality of synthetic genes *in vivo*, propose them as a potential tool for synthetic biology, developmental studies and biomedical applications and challenge the role of the N-terminal part of the endogenous fly protein in transcriptional specificity. This leads to the realization that other roles apart from specificity might be assigned to this long protein sequence, including, but not limited to, transcriptional potency.

The Antp YPWM motif has a dual role in transcriptional activation and repression. To test the full-length and synthetic peptides for their ability to act as transcriptional activators, we used *Tsh* (Fig. 2B) and *grn* (Fig. 5) as *Antp* target genes in the embryo and antennal disc, respectively, and the homeodomain

consensus BS2 binding sites in cultured Drosophila S2 cells (Fig. 6). An absolute requirement for the presence of the tetrapeptide was observed for *Antp* homeotic function in the embryo, larva and adult fly, unlike transactivation experiments where the absence of YPWM did not alter, but rather enhanced, the transcriptional capacity of Antp,

In the case of transcriptional repression, genes known to be, directly or indirectly, downregulated by *Antp* (*Scr* in the embryo, *Salm* and *dan* in the antennal disc) were used as indices for comparative analysis of all constructs. In all cases, a functional repressor motif was found to be necessary for transcriptional repression, since the substitution of the YPWM motif by alanines in the full-length ^{AAAA}*Antp* and synthetic ^{AAAA}*Antp* lines abolished the capacity of the peptides to act as repressors *in vivo*. Substitution of the YPWM motif by WRPW, however, retained to a large extent the repressor activity of both synthetic and full-length peptides and resulted in partial or complete repression of reporter genes. This suggests that other co-repressors interacting with WRPW (including, but not limited to, Gro) might suffice to substitute Pbc-type of co-factors in the contexts studied.

The differential requirements for a functional YPWM motif in Antp support the notion that different rules may apply between *Hox*-mediated transcriptional repression and activation. Thus, the Antp YPWM motif might provide the functional flexibility to recruit coactivators/corepressors, Hox-collaborators, transcriptional machinery factors and/or chromatin remodeling complexes into an Antp-multiprotein complex that allows the fine-tuning of a transcriptional response. Sequence comparisons among different Hox binding preferences collected from studies where there has been substantial evidence for direct Hox binding *in vivo* – reviewed in (Mann et al., 2009) – show that while Pbc-Hox sites are usually found in enhancers of genes activated by a Pbc-Hox complex, rather than repressed, binding sites for Hox factors alone (without any Pbc input) seem to be equally distributed among the regulatory elements of genes activated and repressed by Hox factors. If a functional tetrapeptide is more often a requirement for Antp-mediated activation, these observations do not come as a surprise, but further support the idea that Hox-mediated activation might largely depend on cofactors, while repression might not always require them.

Substitution of the YPWM motif by alanines in full-length and synthetic Antp peptides significantly reduced their homeotic function in the fly (Figs. 1-5). However

these peptides were able to bind Antp binding sites in cultured *Drosophila* cells (Fig. 6 B-C) and trigger transcription of the reporter, suggesting that DNA binding of Antp and the assembly of a functional transcriptional complex can be YPWM- and therefore cofactor-independent. Moreover, in transactivation experiments the YPWM motif might be functioning as a putative repression domain. It is also possible that it alters the kinetics of the different Antp peptides – through interaction with co-factors –and limits their efficiency in binding. This scenario would explain why variants devoid of a YPWM motif exhibited enhanced activity over the wild type protein. Evidence supporting this hypothesis stems from the observations that the presence of both Exd and Bip2 in S2 cells limited the transactivation efficiency of synthetic YPWM-containing peptides (Fig. 7 B-C).

A clear difference in transactivation potency was found between full-length and synthetic AAAA constructs. The former behaved on the average as 50% more potent transactivation factors compared to the wild type Antp protein, while the latter exhibited marginally reduced capacity to activate transcription, as compared to their synthetic ^{YPWM}Antp counterparts.

The seemingly inconsistent behaviour of the ^{AAAA}Antp line in the embryo and larva during development (where it exhibited weak homeotic function), as compared to its behaviour in cultured S2 cells (where it appeared improved over the wild type protein), need not be unusual, since the fusion of consensus binding sites to *hsp70* promoter (Fig. 6 A) facilitates transcriptional regulation also in cases where transcription factor binding alone would not activate transcription in transgenic flies. Such a phenomenon has been observed with *fushi tarazu* (*ftz*) and *engrailed* (*en*) regulatory elements *in vivo* (Vincent et al., 1990). Finally, the presence of the entire N-terminal sequences in the absence of a functional YPWM motif (^{AAAA}Antp) might alternatively facilitate other Hox-cofactor/collaborator interactions that would allow changes in the strength of transcriptional response. Such a scenario would count in favour of the N-terminus participating in potency and would explain why the corresponding synthetic ^{AAAA}Antp constructs fail to exhibit an analogous increase in reporter gene transcription over their YPWM-containing counterparts. One possible N-terminal domain to which potency could be attributed is the poly-glutamine stretch (poly-Q) (present in the N-terminus of Antp and absent in synthetic Antp peptides), which has been originally proposed to function as an activator domain in Bicoid (Janody et al., 2001) and later in Abd-B (Merabet et al., 2003).

Our phenotypic observations in the embryo and adult fly, in which the ^{AAAA}Antp^{WRPW} peptide activated *Tsh* ectopically in the region of the embryonic head, while the synthetic or full-length ^{AAAA}Antp peptides did not, suggest that the fusion of WRPW to the C-terminus of the AAAA-substituted Antp contributes additional effects in transcriptional activation (possibly through repression of a repressor). Similar findings stem from the observation of the tarsal transformation phenotypes in the antennae of adult flies that are indicative of *Antp*-specific function. Substitution of the YPWM motif by alanines, completely abolished *Antp* function in the full-length peptide, while ^{AAAA}Antp^{WRPW} or the synthetic ^{AAAA}Antp constructs, lacking the complete N-terminal sequences, exhibited aristal transformations and/or displayed ectopic leg bristles in the A3 segment, suggesting weak but substantial transformations towards leg identity. A similar behaviour of AAAA-substituted Antp peptides has been reported in flies using the fly full-length ^{AAAA}*Antp*, coinduced with a constitutively active version of *Notch* (N^{act}) (Prince et al., 2008), or the beetle *fushi tarazu* (*ftz*) gene encoding a YPWM-to-AAAA-substituted peptide (Lohr and Pick, 2005). The gain-of-function phenotypes in the antenna are in line with the head involution defects observed in the larval cuticle early in development. Although the AAAA-substituted synthetic peptides overall failed to exhibit any significant phenotype, they still inferred mild malformations in the larval head, including defects in the formation of the mouth hooks and, in these cases, lethality. However, unlike their homeotic effect in the antenna, these lines retained T1 identity in the cuticle, as demonstrated by the presence of prothoracic beards.

The long and short linkers between the YPWM motif and the HD enhance the activation and repression capacity of Antp respectively. The long linker synthetic YPWM construct exhibited stronger *Tsh*-activation capacity in the embryo, resulted in almost complete tarsal transformation of the antenna in adult flies, which very much resembled the effect of the wild type protein, and, finally, enhanced luciferase activity in S2 cells. The short linker construct conferred a more pronounced head involution defect in the larval cuticle and a stronger repression of *Scr* in the embryo. The abundance of Antp transcripts bearing a long (RSQFGKCQ) or a short (RSQF) linker between the YPWM motif and the HD has been found to vary in different developmental stages, with the long linker variant being predominantly expressed in embryonic stages and on the average equal amounts of both variants being detected in larval and pupal stages, as well as in the adult

(Stroeher et al., 1988). Thus, different *Antp* transcripts might play distinct, albeit similar, roles in development. Differences in linker size as a result of alternatively spliced transcripts has been found also for Ubx (O'Connor et al., 1988) and participation of the linker region in transcriptional control has been reported for AbdA, in which mutations in the linker decreased the activation capacity of the transcription factor on the *wingless* (*wg*) enhancer (Merabet et al., 2003). Our results support these findings: the short-linker construct appeared to be a more powerful repressor, and the long-linker construct a more powerful activator, of *Antp* target genes, suggesting a role of the linker size and/or sequence in transcriptional potency/specificity *in vivo*. The distance between the Antp YPWM motif and its HD might also affect DNA-binding affinity of the HD N-terminal arm, as previously reported with N-terminal arm-depleted Antp (Qian et al., 1994b). Alternatively, linker regions of Antp may also interact with additional cofactors or proteins in a DNA-binding-dependent or independent manner (Mann et al., 2009; Merabet et al., 2003).

Transcriptional activation of synthetic Antp peptides is altered in the presence of Exd or Bip2 in a YPWM-dependent fashion. Finally we found that synthetic Antp peptides were able to activate transcription from the BS2 sites. Therefore, we investigated the transactivation behaviour of synthetic peptides in the presence of Exd (Passner et al., 1999; Piper et al., 1999; Ryoo and Mann, 1999) and Bip2 (Prince et al., 2008), co-factors that interact with Antp via its YPWM motif. In transactivation experiments the efficiency of both synthetic ^{YPWM}Antp peptides was altered in the presence of both Exd and Bip2 and a clear improvement in transcriptional potency was found upon coexpression of ^{YPWM,SL}Antp and Exd, further supporting the notion that the linker region contributes significantly to the correct positioning of transcription factor and co-factor, such that DNA-binding and interactions between the two peptides and with the transcriptional machinery are facilitated. Unlike the YPWM peptides, the AAAA peptides, as expected, did not interact with Exd or Bip2, since a functional YPWM motif is necessary for these interactions.

A C-terminal portion of Antp, containing the HD and YPWM motif, suffices for homeotic function in *Drosophila*. As in the case of synthetic *Scr* genes (Papadopoulos et al., 2010), our results challenge the role of the Antp long N-terminal sequence in transcriptional specificity and support previous findings of us and others (Gibson et al., 1990; Shen et al., 2004b; Zhao et al., 1996) that the Hox

N-terminus plays a role in transcriptional potency, possibly by regulating the amount of transcripts produced and/or the persistence of transcription in time, rather than participating in decisions on which effector genes might be regulated by which Hox factor at which time during development. The aforementioned poly-Q stretches could qualify as a candidate domain to which potency might be attributed. Another conserved motif is the SSYF (TSYF in Antp), decorating the N-terminus of many homeoproteins (Tour et al., 2005) and described to participate in transcriptional potency in the fly Scr (Tour et al., 2005) and the mouse Hoxa5 (Zhao et al., 1996). However, if this is the only function of the N-terminus, how can the C-terminal portion of the protein alone confer the required specificity and differentially regulate putative target genes in a context-specific manner? The plausible explanations remain scarce, since cofactors known to date to interact with Antp are limited (Mann et al., 2009) and the specificity of its HD, also assayed *in vitro* (Affolter et al., 1990; Muller et al., 1988) cannot account for target selection *in vivo*. However, the linker region between the HD and the YPWM motif, differentially spliced in the transcripts of at least some *Hox* genes (*Antp*, *Ubx*), could account for increased specificity of target selection *in vivo*. Recent evidence shows that the regions flanking the Ubx YPWM motif (alternatively spliced in Ubx isoforms) account for specificity of Ubx *in vivo* (Liu et al., 2009; Reed et al.). Furthermore, decisions on whether certain Hox factors will function as activators versus repressors in a specific context might as well be owed to the abundance of their isoforms, rather than the regulatory sequences to which they may bind. The finding that the same enhancer element upstream of the murine *Six2* gene is activated by Hox11 and repressed by Hoxa2 (Yallowitz et al., 2009) substantiates the aforementioned hypothesis.

The current study has also increased our understanding about synthetic genes. Hypothesizing that protein function is a sum of the functions of its protein domains, the latter being oriented in a precise three-dimensional conformation that allows normal interactions to take place in the correct way, maybe we are not so far from attempting to engineer and possibly predict the functions of novel synthetic peptides on the basis of the structural architecture of their protein domains. Inversely, functions of proteins might be limited down to fractions of their sequences, with each fraction retaining only a subset of these functions. So far, we have successfully reconstructed to a great extent the function of two *Hox* genes, *Antp* and *Scr*, using their synthetic peptide-versions, and demonstrated that they act

predictably *in vivo*. If biomedically relevant genes also fell in this category, the development and usage of their corresponding synthetic peptides might bear advantages over the full-length proteins for application *in vivo*, due to their smaller size and easier-to-understand molecular properties.

Materials and Methods

Generation of constructs and fly transgenesis. Plasmids were generated using standard procedures and fly transgenesis was performed as described (Spradling and Rubin, 1982). All transgenic UAS lines used are shown in Fig. 1 Q. Further lines used here were: *ptc-Gal4*, *nullo-Gal4* (Gehring et al., 2009), *Dll-Gal4*(*MD23*)/*CyO; Dr*/*TM6B* and *Dll-Gal4(MD713)/Dll-Gal4(MD713)* (Calleja et al., 1996), *CyO*/*Sp*; *grn-lacZ*/*TM6B* (Grieder et al., 2009; Wagner-Bernholz et al., 1991), *CyOSalm-lacZ/Sp; Dr/TM6B* (Wagner-Bernholz et al., 1991), *Dll-Gal4*(*MD23*)/*CyO*; *grn-lacZ*/*TM6B*.

Immunohistochemistry on embryos. Embryos of the cross of *ptc*-Gal4 (or *nullo*-Gal4) to all different UAS-*Antp* lines were collected for 2 h at 25° C, after a precollection of 2 h at 25° C to allow synchronized development, then aged at 25° C for 6.5 h, dechorionated using 3-4% chlorax, washed with 120 mM NaCl, 0.03% TritonX-100, dried and fixed under continuous, vigorous rotation for 20 min at room temperature (RT) in 1 mL 1:1 solution of n-heptane/embryo fixation solution (embryo fix: 2X PBS, 0.02% Tween^{20}, 13.8% EM-grade paraformaldehyde (PFA) (Electron Microscopy Science, PA, USA) freshly used every 10 days). After removal of the lower phase of the solution (n-heptane) and addition of 700 µL methanol, embryos were vigorously shaken for 5 min, until the sunk to the bottom, wahed 3 time for 20 min at RT with methanol, rehydrated in a 1:1 solution of methanol/PBT (PBS + 0.1% TritonX-100), washed 3 time in PBT, blocked for 20 min at 4° C with 10% fraction V BSA in PBT, washed once for 10 min in PBT and incubated with primary antibodies in the appropriate dilutions in PBT, containing 0.02% NaN_3 and 5% normal goat serum (NGS) (Jackson ImmunoResearch) overnight at 4° C under continuous slow rotation. Subsequently, embryos were rinsed once in PBT, washes 3 times for 20 min with PBT at RT and incubated with secondary antibodies for 4 h at RT under continuous slow rotation, then rinsed once and washed 3 times with PBT at RT, equilibrated in PBS. After removal of PBS, embryos were mounted in in Vectashield

(Vector Laboratories). Images were obtained within 24 h using a Leica SP5 Confocal setup. Antibodies were used in the following dilutions: Tsh (rabbit) 1:2000 (Wu and Cohen, 2000); Antp 4C3 (mouse) 1:250 (Developmental Studies Hybridoma Bank, University of Iowa); Scr (rabbit) 1:200 (LeMotte et al., 1989); GFP (mouse) 1:250 (Invitrogen) and GFP (rabbit) 1:500 (Invitrogen).

Immunohistochemistry on imaginal discs. Third instar wandering larvae were dissected at RT in Grace's medium (Gibco) and fixed for 10 min in Grace's medium containing 5.3% PFA (see above). After twice washed for 15 min in PBT and blocked for at least 30 min in PBT containing 5% NGS and 20 µg/mL NaN3, samples were incubated with primary antibodies in blocking solution overnight. Subsequently, samples were washed 3 times for 20 min in PBT and secondary antibodies were applied in PBS for 4 h at RT. After 3 washes for 20 min with PBT and equilibration in PBS, discs were dissected from the fixed carcasses and mounted in Vectashield. Antibodies were used in the following dilutions: Salm (rabbit) 1:25 (Kuhnlein et al., 1994); β-gal (mouse) 1:500 (Promega); dan (rat) 1:300 (Emerald et al., 2003);

Preparation of embryonic cuticle. The same procedure as the one for antibody staining was followed, but embryos were rather allowed to develop for 1.5-2 days after egg-collection, this resulting in the separation of the survivors, which crawled into the love yeast, from non-survivors. Then embryos were subjected to the same treatments as in the staining protocol, omitting only the fixation step. After washes in methanol, the latter was removed completely and embryos were mounted in a 1:1 Hoyer's/90% lactic acid solution. Then they were mounted on microscopic slides, covered with coverslips and subjected to digestion for 24 h at 58° C, flattening for 24 h at RT and used for observation.

Preparation of adult cuticle. Escapers or pharate adult flies from the cross of *Dll*-Gal4 to the various UAS-*Antp* lines were collected and fixed in 30% glycerol in ethanol. Then, fly parts were dissected from the flies in the same solution, transferred to microscopic slides into droplets of Faure's medium, covered with coverslips, flattened and used for imaging.

Transactivation in Drosophila S2 cells. For transfection assays on Schneider cells (S2), the sequences encoding Antp full-length or synthetic genes were cloning into pNPAC expression vector that contains an Actin5C promoter (Krasnow et al., 1989). Substitution of the YPWM to WRPW was performed using

the QuickChange® XL Site-Directed Mutagenesis kit (Stratagene). pGLH11 contains 11 BS2 binding sites obtained from pF11CAT (Furukubo-Tokunaga et al., 1992). pcopiaβgal reporter plasmids were previously described (Furukubo-Tokunaga et al., 1992). S2 cells were maintained at 25°C in Schneider Medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum heat inactivated (GIBCO- Invitrogen). Cells were seeded into 6-well plates and were transfected at approximately 80% confluency with 0.5 µg of *Antp* full-length or synthetic genes, 0.5 µg of pPACBip2 or Exd, 1 µg of pGLH11 and 0.5 µg of pcopiaβgal as an internal control for transfection efficiency. The amount of DNA in each well was kept constant by addition of appropriate amount of the parental expression plasmid pNPAC. For all transfections the Cellfectin Reagent was used as recommended by the manufacturer (Invitrogen). Transfection reaction was left on the cells for 5 hours and 72 hours after transfection, cells were harvested in 1X Reporter lysis buffer (Promega). β-Galactosidase activity was assayed essentially as previously described (Furukubo-Tokunaga et al., 1992). Luciferase activity was measured as relative luminescence units in an Optocomp I luminometer (MGM Instruments) using a luciferase assay system (Promega) and normalized against β-galactosidase activity control. Experiments were performed by triplicate and repeated a minimum of three times.

Acknowledgements

This work was supported by the Kantons of Basel-Stadt and Basel-Landschaft, a grant from the Swiss National Science Foundation, the European Network of Excellence “Cells into Organs” and the Consejo Nacional de Ciencia y Tecnología de México Beca Nacional Número 46659.

References

Affolter, M., Percival-Smith, A., Muller, M., Leupin, W. and Gehring, W. J. (1990). DNA binding properties of the purified Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **87**, 4093-7.

Andrew, D. J., Horner, M. A., Petitt, M. G., Smolik, S. M. and Scott, M. P. (1994). Setting limits on homeotic gene function: restraint of Sex combs reduced activity by teashirt and other homeotic genes. *EMBO J* **13**, 1132-44.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-15.

Calleja, M., Moreno, E., Pelaz, S. and Morata, G. (1996). Visualization of gene expression in living adult Drosophila. *Science* **274**, 252-5.
Emerald, B. S., Curtiss, J., Mlodzik, M. and Cohen, S. M. (2003). Distal antenna and distal antenna related encode nuclear proteins containing pipsqueak motifs involved in antenna development in Drosophila. *Development* **130**, 1171-80.
Fisher, A. L., Ohsako, S. and Caudy, M. (1996). The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. *Mol Cell Biol* **16**, 2670-7.
Furukubo-Tokunaga, K., Flister, S. and Gehring, W. J. (1993). Functional specificity of the Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **90**, 6360-4.
Furukubo-Tokunaga, K., Muller, M., Affolter, M., Pick, L., Kloter, U. and Gehring, W. J. (1992). In vivo analysis of the helix-turn-helix motif of the fushi tarazu homeo domain of Drosophila melanogaster. *Genes Dev* **6**, 1082-96.
Galant, R., Walsh, C. M. and Carroll, S. B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* **129**, 3115-26.
Gangloff, Y. G., Pointud, J. C., Thuault, S., Carre, L., Romier, C., Muratoglu, S., Brand, M., Tora, L., Couderc, J. L. and Davidson, I. (2001). The TFIID components human TAF(II)140 and Drosophila BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol Cell Biol* **21**, 5109-21.
Gehring, W. J. (1987). Homeo boxes in the study of development. *Science* **236**, 1245-52.
Gehring, W. J., Kloter, U. and Suga, H. (2009). Evolution of the Hox gene complex from an evolutionary ground state. *Curr Top Dev Biol* **88**, 35-61.
Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994). Homeodomain-DNA recognition. *Cell* **78**, 211-23.
Gibson, G., Schier, A., LeMotte, P. and Gehring, W. J. (1990). The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**, 1087-103.
Grieder, N. C., Morata, G., Affolter, M. and Gehring, W. J. (2009). Spalt major controls the development of the notum and of wing hinge primordia of the Drosophila melanogaster wing imaginal disc. *Dev Biol* **329**, 315-26.
Hueber, S. D. and Lohmann, I. (2008). Shaping segments: Hox gene function in the genomic age. *Bioessays* **30**, 965-79.
Janody, F., Sturny, R., Schaeffer, V., Azou, Y. and Dostatni, N. (2001). Two distinct domains of Bicoid mediate its transcriptional downregulation by the Torso pathway. *Development* **128**, 2281-90.
Johnson, F. B., Parker, E. and Krasnow, M. A. (1995). Extradenticle protein is a selective cofactor for the Drosophila homeotics: role of the homeodomain and YPWM amino acid motif in the interaction. *Proc Natl Acad Sci U S A* **92**, 739-43.
Joshi, R., Passner, J. M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A., Jacob, V., Aggarwal, A. K., Honig, B. and Mann, R. S. (2007). Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* **131**, 530-43.
Kmita-Cunisse, M., Loosli, F., Bierne, J. and Gehring, W. J. (1998). Homeobox genes in the ribbonworm Lineus sanguineus: evolutionary implications. *Proc Natl Acad Sci U S A* **95**, 3030-5.

Kuhnlein, R. P., Frommer, G., Friedrich, M., Gonzalez-Gaitan, M., Weber, A., Wagner-Bernholz, J. F., Gehring, W. J., Jackle, H. and Schuh, R. (1994). spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. *EMBO J* **13**, 168-79.

LeMotte, P. K., Kuroiwa, A., Fessler, L. I. and Gehring, W. J. (1989). The homeotic gene Sex Combs Reduced of Drosophila: gene structure and embryonic expression. *EMBO J* **8**, 219-27.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565-70.

Liu, Y., Matthews, K. S. and Bondos, S. E. (2009). Internal regulatory interactions determine DNA binding specificity by a Hox transcription factor. *J Mol Biol* **390**, 760-74.

Lohr, U. and Pick, L. (2005). Cofactor-interaction motifs and the cooption of a homeotic Hox protein into the segmentation pathway of Drosophila melanogaster. *Curr Biol* **15**, 643-9.

Mann, R. S. and Chan, S. K. (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet* **12**, 258-62.

Mann, R. S., Lelli, K. M. and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. *Curr Top Dev Biol* **88**, 63-101.

McCormick, A., Core, N., Kerridge, S. and Scott, M. P. (1995). Homeotic response elements are tightly linked to tissue-specific elements in a transcriptional enhancer of the teashirt gene. *Development* **121**, 2799-812.

McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.

McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. and Gehring, W. J. (1984). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* **308**, 428-33.

Merabet, S., Kambris, Z., Capovilla, M., Berenger, H., Pradel, J. and Graba, Y. (2003). The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Dev Cell* **4**, 761-8.

Muller, M., Affolter, M., Leupin, W., Otting, G., Wuthrich, K. and Gehring, W. J. (1988). Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J* **7**, 4299-304.

O'Connor, M. B., Binari, R., Perkins, L. A. and Bender, W. (1988). Alternative RNA products from the Ultrabithorax domain of the bithorax complex. *EMBO J* **7**, 435-45.

Papadopoulos, D. K., Vukojevic, V., Adachi, Y., Terenius, L., Rigler, R. and Gehring, W. J. (2010). Function and specificity of synthetic Hox transcription factors in vivo. *Proc Natl Acad Sci U S A* **107**, 4087-92.

Paroush, Z., Finley, R. L., Jr., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R. and Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**, 805-15.

Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S. and Aggarwal, A. K. (1999). Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* **397**, 714-9.

Pearson, J. C., Lemons, D. and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. *Nat Rev Genet* **6**, 893-904.

Peifer, M. and Wieschaus, E. (1990). Mutations in the Drosophila gene extradenticle affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev* **4**, 1209-23.
Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L. and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* **96**, 587-97.
Prince, F., Katsuyama, T., Oshima, Y., Plaza, S., Resendez-Perez, D., Berry, M., Kurata, S. and Gehring, W. J. (2008). The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* **135**, 1669-79.
Qian, Y. Q., Resendez-Perez, D., Gehring, W. J. and Wuthrich, K. (1994). The des(1-6)antennapedia homeodomain: comparison of the NMR solution structure and the DNA-binding affinity with the intact Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **91**, 4091-5.
Rawat, V. P. S., Thoene, S., Naidu, V. M., Arseni, N., Heilmeier, B., Metzeler, K., Petropoulos, K., Deshpande, A., Quintanilla-Martinez, L., Bohlander, S. K. et al. (2008). Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood* **111**, 309-319.
Reed, H. C., Hoare, T., Thomsen, S., Weaver, T. A., White, R. A., Akam, M. and Alonso, C. R. Alternative Splicing Modulates Ubx Protein Function in Drosophila melanogaster. *Genetics*.
Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S. and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* **461**, 1248-53.
Ryoo, H. D. and Mann, R. S. (1999). The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* **13**, 1704-16.
Shen, W. F., Chrobak, D., Krishnan, K., Lawrence, H. J. and Largman, C. (2004). HOXB6 protein is bound to CREB-binding protein and represses globin expression in a DNA binding-dependent, PBX interaction-independent process. *Journal of Biological Chemistry* **279**, 39895-39904.
Shen, W. F., Rozenfeld, S., Lawrence, H. J. and Largman, C. (1997). The Abd-B-like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target. *J Biol Chem* **272**, 8198-206.
Spradling, A. C. and Rubin, G. M. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. *Science* **218**, 341-7.
Stroeher, V. L., Gaiser, J. C. and Garber, R. L. (1988). Alternative RNA splicing that is spatially regulated: generation of transcripts from the Antennapedia gene of Drosophila melanogaster with different protein-coding regions. *Mol Cell Biol* **8**, 4143-54.
Tolkunova, E. N., Fujioka, M., Kobayashi, M., Deka, D. and Jaynes, J. B. (1998). Two distinct types of repression domain in engrailed: one interacts with the groucho corepressor and is preferentially active on integrated target genes. *Mol Cell Biol* **18**, 2804-14.
Tour, E., Hittinger, C. T. and McGinnis, W. (2005). Evolutionarily conserved domains required for activation and repression functions of the Drosophila Hox protein Ultrabithorax. *Development* **132**, 5271-81.
Vincent, J. P., Kassis, J. A. and O'Farrell, P. H. (1990). A synthetic homeodomain binding site acts as a cell type specific, promoter specific enhancer in Drosophila embryos. *EMBO J* **9**, 2573-8.

Vukojevic, V., Papadopoulos, D. K., Terenius, L., Gehring, W. J. and Rigler, R. (2010). Quantitative study of synthetic Hox transcription factor-DNA interactions in live cells. *Proc Natl Acad Sci U S A* **107**, 4093-8.
Wagner-Bernholz, J. T., Wilson, C., Gibson, G., Schuh, R. and Gehring, W. J. (1991). Identification of target genes of the homeotic gene Antennapedia by enhancer detection. *Genes Dev* **5**, 2467-80.
Wu, J. and Cohen, S. M. (2000). Proximal distal axis formation in the Drosophila leg: distinct functions of teashirt and homothorax in the proximal leg. *Mech Dev* **94**, 47-56.
Yallowitz, A. R., Gong, K. Q., Swinehart, I. T., Nelson, L. T. and Wellik, D. M. (2009). Non-homeodomain regions of Hox proteins mediate activation versus repression of Six2 via a single enhancer site in vivo. *Dev Biol* **335**, 156-65.
Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. and Scott, M. P. (1993). Ectopic expression and function of the Antp and Scr homeotic genes: the N terminus of the homeodomain is critical to functional specificity. *Development* **118**, 339-52.
Zhao, J. J., Lazzarini, R. A. and Pick, L. (1996). Functional dissection of the mouse Hox-a5 gene. *EMBO J* **15**, 1313-22.

Figures

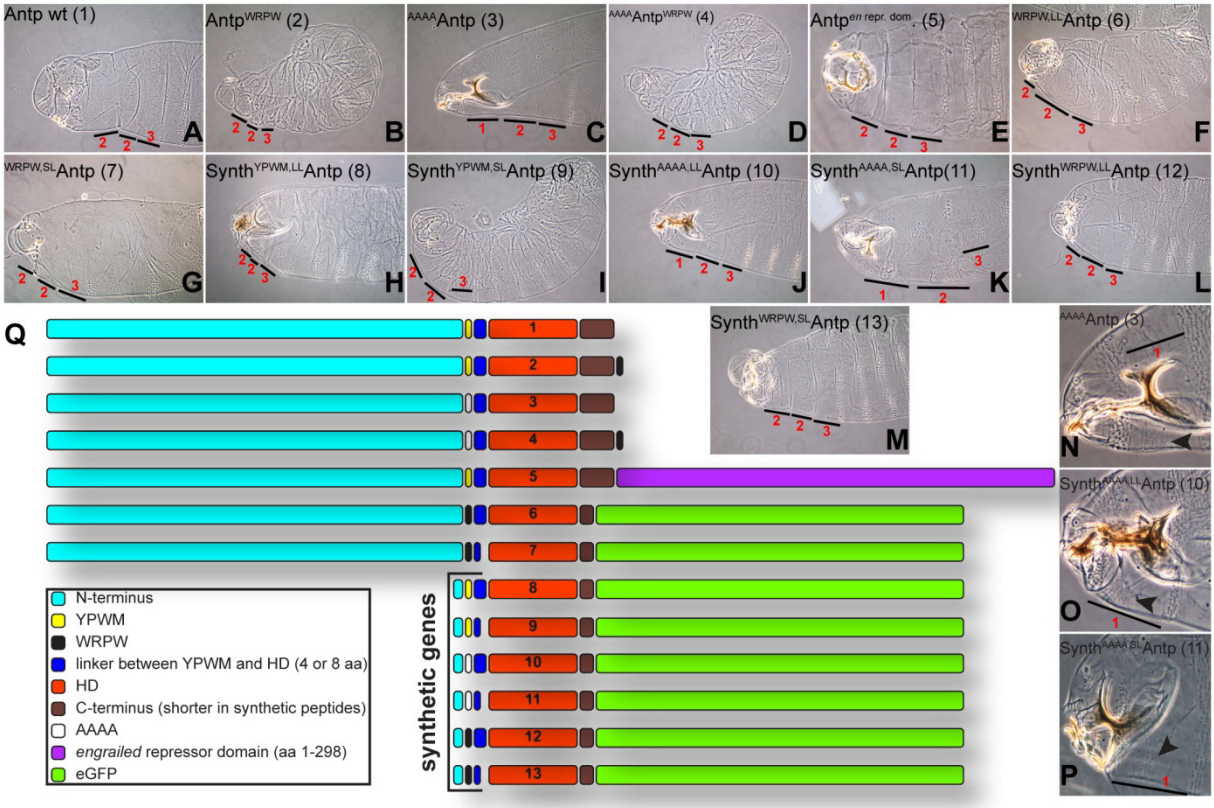

Figure 1. Full-Length and Synthetic *Antp* Genes Cause Homeotic Transformations and Head Involution Defects in the Larval Cuticle. (A-M) Comparative analysis of head involution defects generated by constitutive expression of full-length and synthetic *Antp* genes, using the nullo-Gal4 driver line. Transformation of the head structures towards a mesothoracic identity causes head involution defects and transformation of the prothoracic segment (T1), normally specified by *Scr* towards a mesothoracic segment (T2), normally specified by *Antp*, results in loss of the prothoracic beards that characterize T1 identity. All *Antp* constructs are coded by numbers (1-13) and represented schematically below. The size of each thoracic segment in all embryos is indicated by lines spanning the segments. The numbers (1-3) below or above these lines show the corresponding segment identity: prothoracic (1), mesothoracic (2), or metathoracic (3). In the lines where homeotic transformations are observed, the identity of the corresponding segments is indicated by the corresponding number. (N-P) Magnification of anterior structures of lines expressing AAAA-substituted Antp peptides, where the presence of prothoracic beards (arrowheads) indicates absence of transformation. Note the mild head defects in synthetic *Antp*^{AAAA} lines (O-P), shown by the abnormal formation of the mouth hooks. (Q) Schematic representation of UAS constructs generated (drawn to scale). The numbers in the constructs represent the different genotypes shown above.

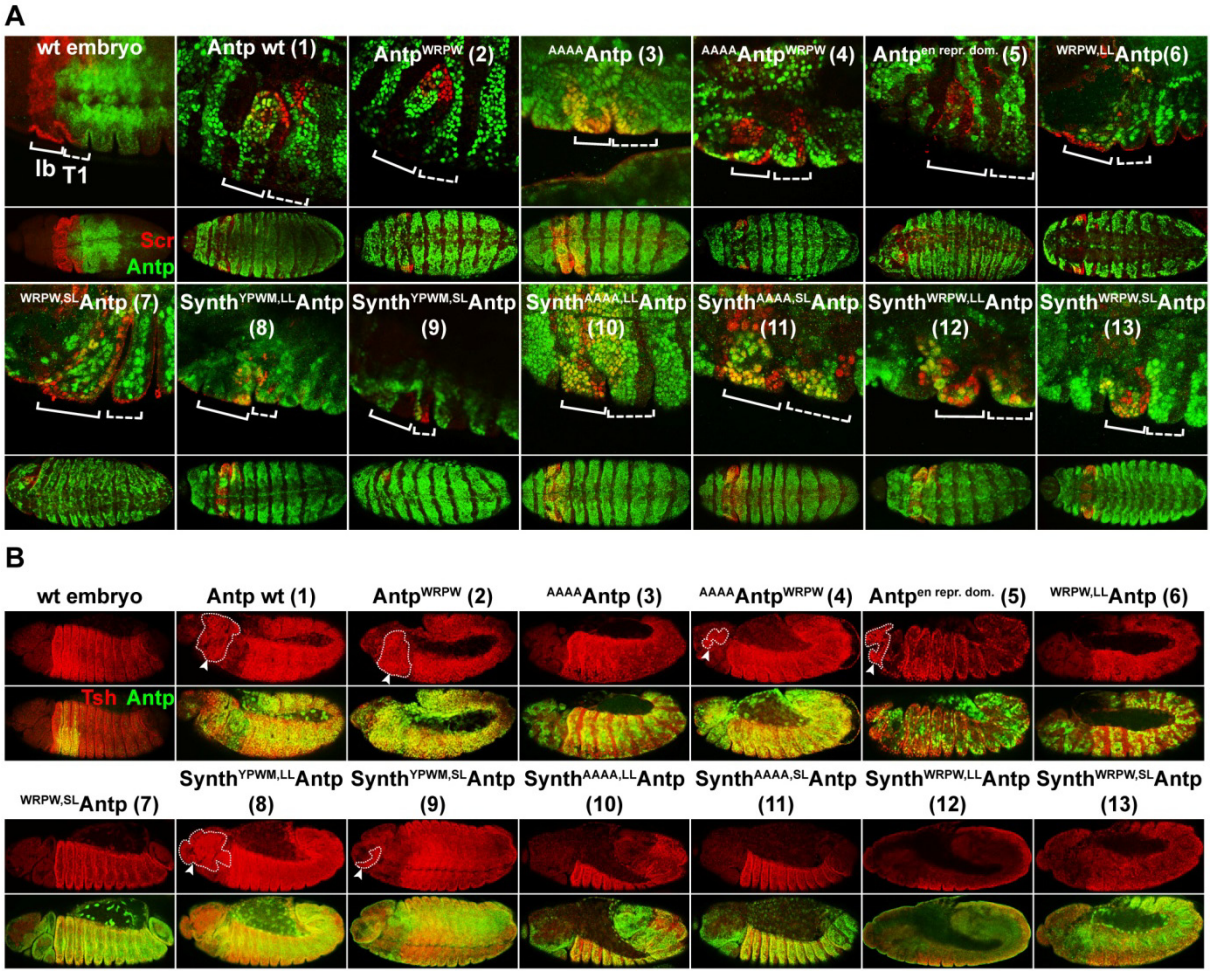

Figure 2. Repression and Activation of Antp Target Genes in the Embryo Mediated by Full-Length and Synthetic *Antp* Genes. (A) Repression of *Scr* in the labial (lb, indicated in wild type embryo, solid brackets) and prothoracic (T1, indicated in wild type embryo, dashed brackets) segments mediated by full-length and synthetic *Antp* constructs. For each line a close-up of the region of repression and, below it, an overview of an entire embryo, are shown. Ectopic Antp is depicted in green and Scr in red. Note the overlay of both gene products in the AAAA-substituted lines, indicated by the yellow colour. Staining in the wild type (wt) embryo is for the endogenous genes. (B) Ectopic activation of *Tsh* in the embryonic head segments mediated by full-length and synthetic *Antp* genes. Tsh is shown in red and ectopic Antp in green. The upper images for each line depict the distribution of Tsh and the lower overlays with ectopic Antp. The domain of ectopic *Tsh* expression is outlined by closed dashed lines and shown by arrowheads. Staining in the wild type (wt) embryo is for the endogenous genes. The numbers in brackets (in A and B) correspond to the numbered constructs in Fig. 1 Q. All UAS lines were induced using a *ptc*-Gal4 driver line.

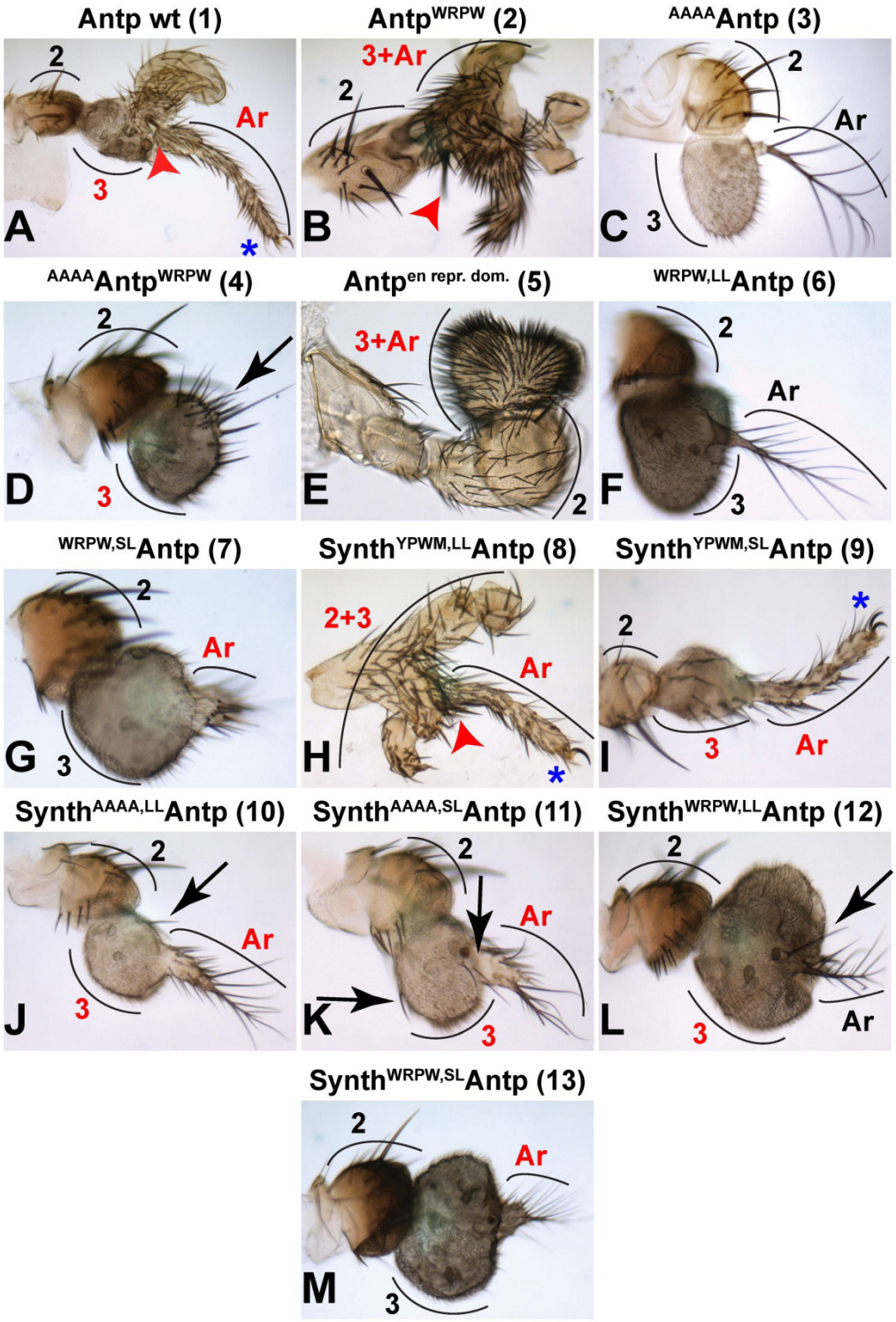

Figure 3. Tarsal Transformations Caused by Full-length and Synthetic Antp Peptides. (A-M) Tarsal transformations induced by synthetic and full-length Antp lines, using a *Dll*-Gal4 driver**.** The presence of the YPWM motif in Antp full-length and synthetic peptides greatly enhanced the transformation, but its absence did not completely abolish it in synthetic peptides. Note the leg bristles (black arrows in D and J-L) that are indicative of a weak transformation of the third antennal segment (A3) to a tarsal fate, as well as the presence of apical/preapical bristles (red arrowheads in A, B and H), which indicate a mesothoracic leg (T2) identity, normally specified by *Antp***.** Antennal segments are indicated by curved lines. The numbers next to them correspond to different segments (2, second antennal segment; 3, third antennal segment; Ar, arista). Transformed segments are indicated using red numbers. Note the presence of tarsal claws (blue asterisks in A, H and I), identified in the cases of strong transformations. Numbers next to the name of each Antp line are explained in Fig. 1 Q.

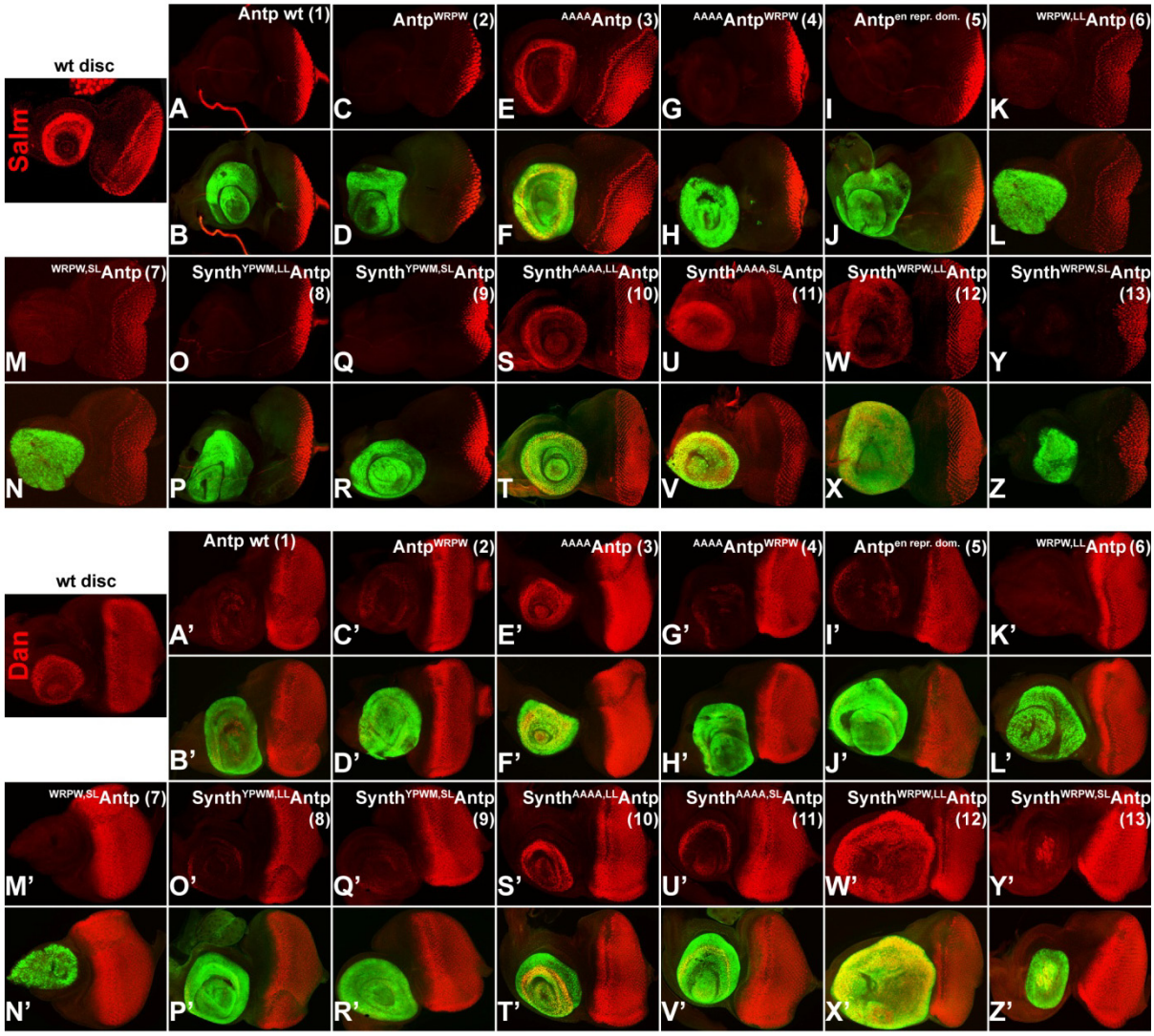

Figure 4. Repression of Genes Required for Normal Antennal Development by Ectopic Expression of Full-Length and Synthetic *Antp* Genes in the Antennal Disc. (A-Z) Repression of *Salm* in the antennal disc mediated by full-length and synthetic *Antp* genes. All peptides featuring a functional YPWM motif or a WRPW repressor motif repressed *Salm*, except for the AAAA-substituted constructs (E-F and S-V) and discs expressing the synthetic *Antp*$^{WRPW,\ long\ linker}$ (W-X). Note the enlargement of the antennal portion in this case. These discs give rise to an enlarged A3 segment in adult antennae (Fig. 3 L). (A'-Z') Repression of *dan* in the antennal disc mediated by full-length and synthetic *Antp* genes. The same behaviour as in the repression of *Salm* was observed.

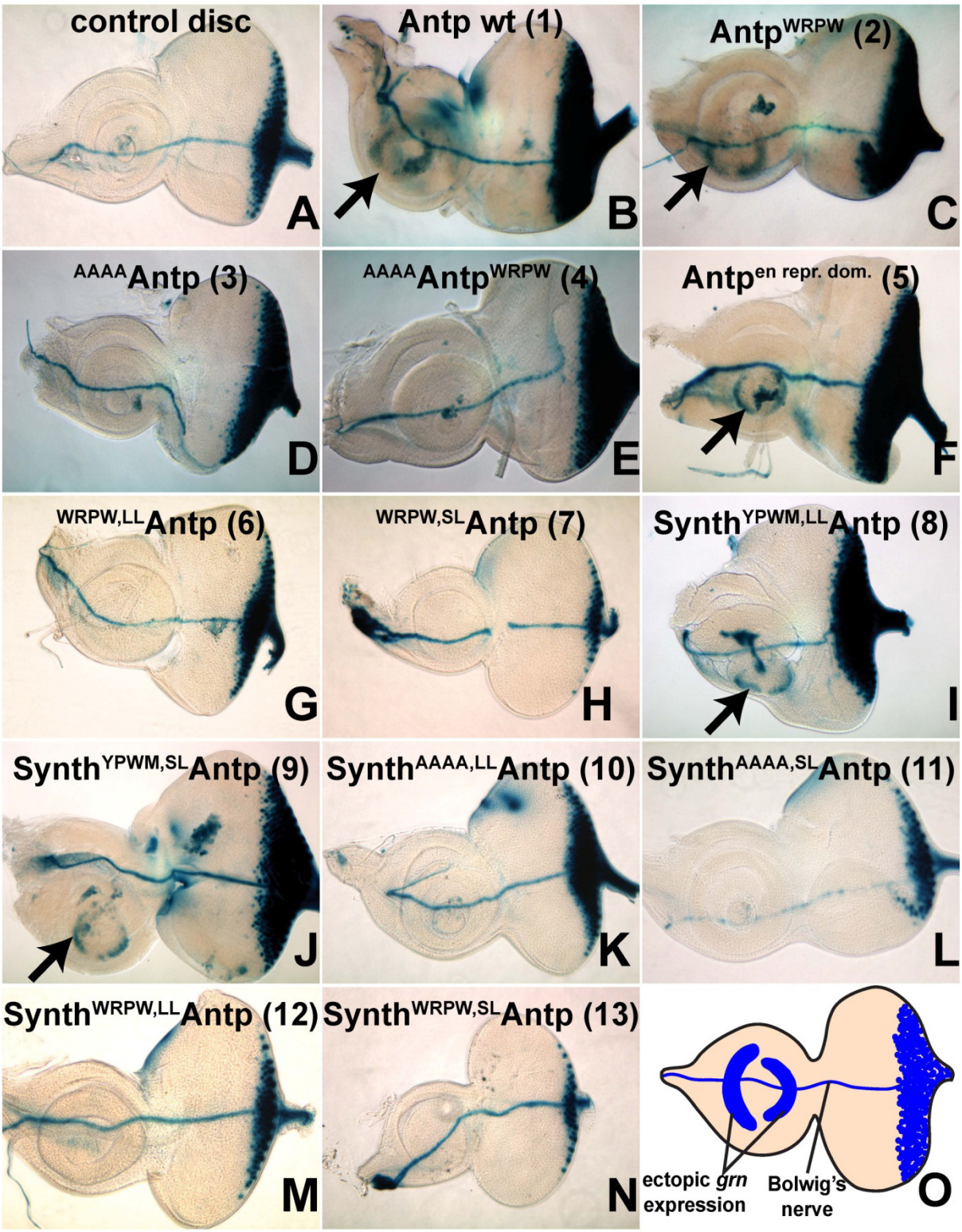

Figure 5. Ectopic Activation of Leg-Specific Genes by Misexpression of Full-Length and Synthetic *Antp* Genes in the Antennal Disc. (A-L) Ectopic activation of *grn* in the antennal disc, mediated by synthetic and full-length Antp constructs. A requirement for a functional YPWM motif was observed. Arrows show the ectopic expression of β-galactosidase in the antennal portion (A-B, E and H-I). (O) Schematic representation of an eye-antennal disc expressing ectopically *grn* upon *Antp* gain-of-function. All UAS constructs (A-L) were induced using a *Dll*-Gal4 driver. The numbers in parentheses correspond to the constructs schematically represented in Fig. 1 Q.

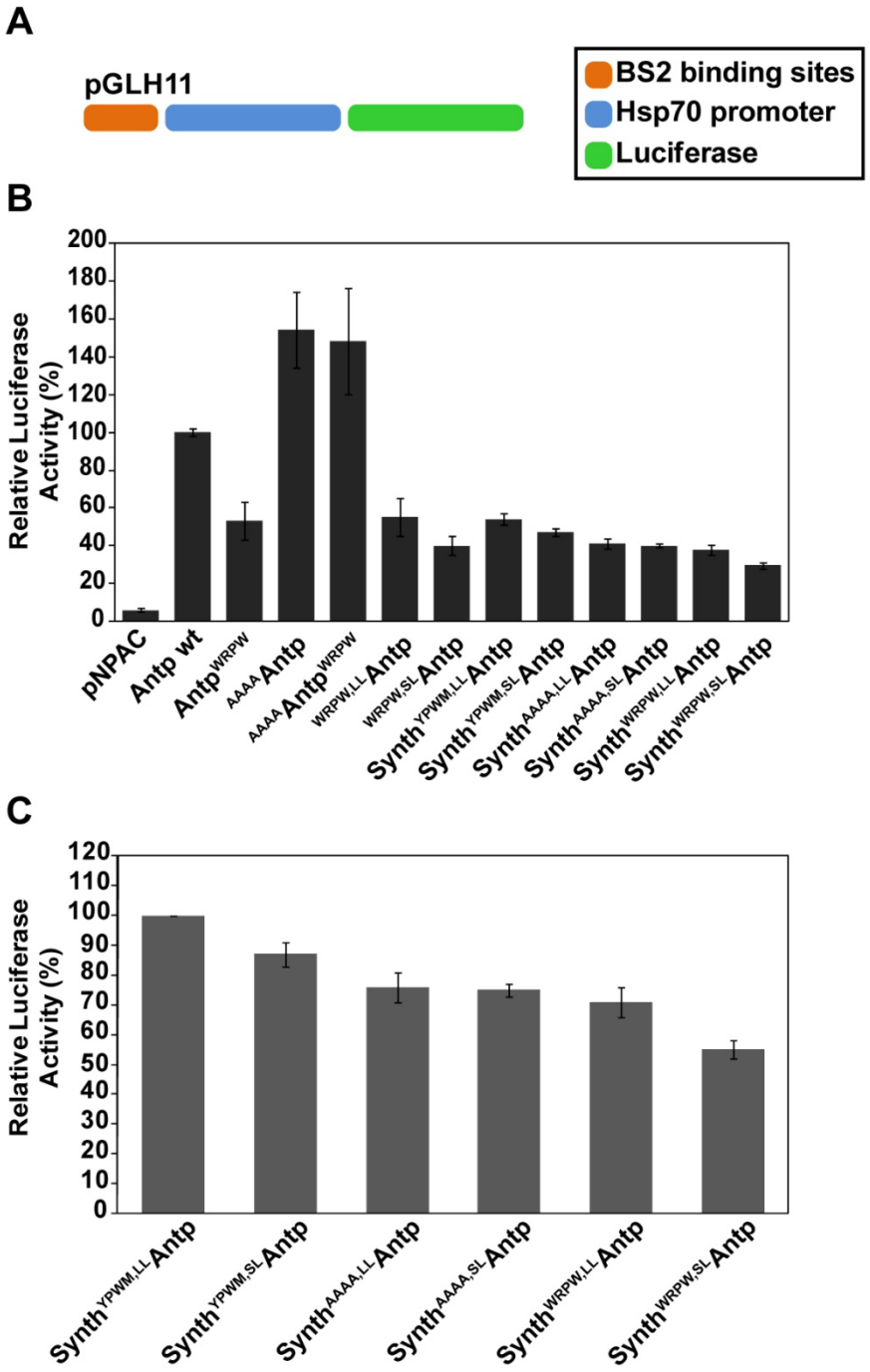

Figure 6. Synthetic and Full-Length Antp Peptides Activate Transcription from the BS2 Site. (A) Schematic representation of the luciferase reporter construct used in transfection assays. It contains a minimal hsp70 promoter and eleven tandem copies of homeodomain consensus binding sites (BS2). (B) Both Antp full-length and synthetic peptides activated transcription, but the synthetic peptides exhibited on the average 50% transactivation activity, as compared to the wild type protein. Reduction of reporter gene expression was observed when WRPW was fused to the C-terminus of Antp. Opposite effects were observed among constructs bearing substitutions of the YPWM motif. While the ^{AAAA}Antp exhibited a clear increment in reporter gene expression, ^{WRPW}Antp behaved on the average as a 50% weaker activator than the wild type protein. (C) Analysis of the differential transactivation efficiency among synthetic Antp constructs. Differences in between pairs of long and short linker variants suggested a role of the linker residues in transcriptional activity. The long linker variants behaved as more potent activators. Luciferase activity was normalized for transfection efficiency using a β-galactosidase control plasmid and transactivation activity of wild type Antp or Synth^{YPWM,LL}Antp was defined as 100%. Each bar represents the average of three independent experiments (± standard deviation) with each sample repeated by triplicate.

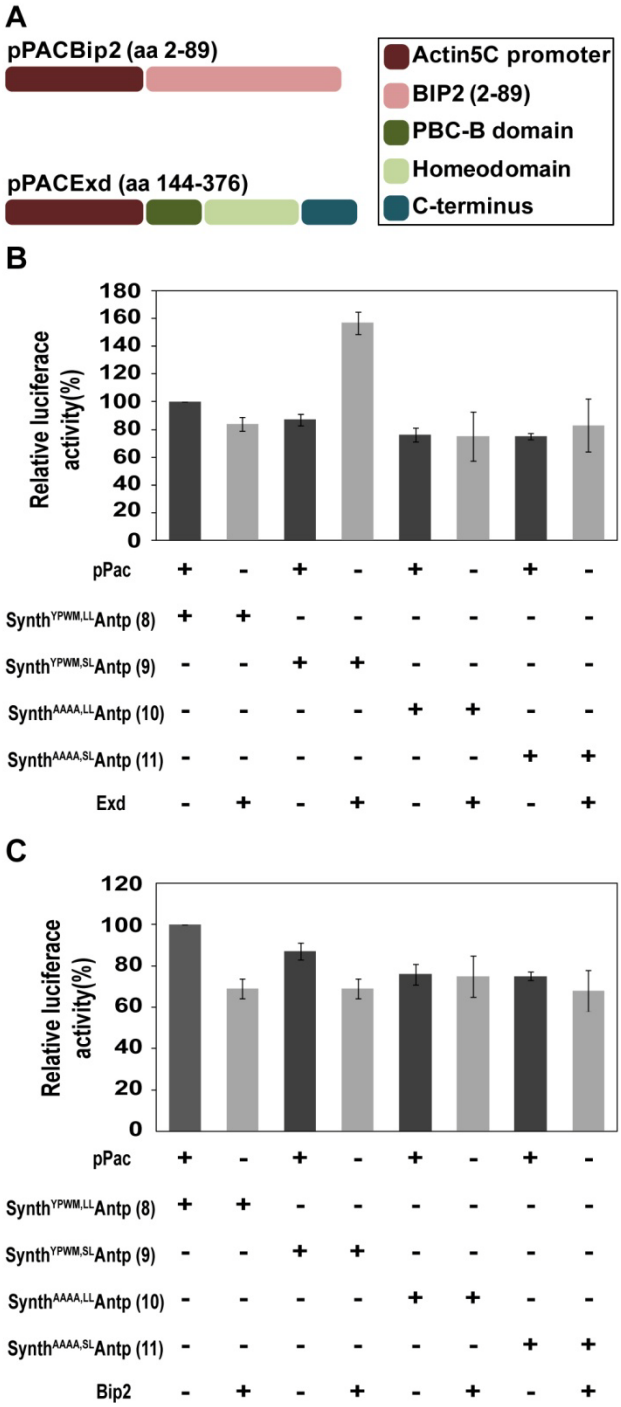

Figure 7. Exd and Bip2 Cofactors Change the Transactivation Properties of Synthetic Antp Peptides in a YPWM-Dependent Way. (A) Transient co-transfection assays of Schneider S2 cells were performed with synthetic *Antp* plasmids (dark gray) and cofactor expression plasmids (light gray) containing sequences encoding Exd (amino acids 144-376, including its PBC-B domain, homeodomain and C-terminal portion) and Bip2 (amino acids 2-89) driven by Actin5C promoter. (B) Co-expression of Exd with Synth^{YPWM,LL}Antp resulted in about 20% repression of the luciferase activity in S2 cells, in contrast to Synth^{YPWM,SL}Antp which increased transcriptional activation to about 70%. No repression or activation was observed upon co-expression of Synth^{AAAA}Antp with Exd. (C) Co-expression of Bip2 and Synth^{YPWM,LL}Antp or Synth^{YPWM,SL}Antp peptides in S2 cells showed 30% repression in transcriptional activity, whereas the Synth^{AAAA}Antp long and short linker genes conferred no transcriptional repression in presence of Bip2. All co-transfection assays included a pcopia-β-gal plasmid as an internal control of transfection efficiency. Activity of luciferase reporter with Synth^{YPWM,LL}Antp was defined as 100%. Bars represent an average of four independent assays (± standard deviation).

Gain-of-function assay	Antp construct												
	Antp wt (1)	Antp^{WRPW} (2)	^{AAAA}Antp (3)	^{AAAA}Antp ^{WRPW} (4)	Antp$^{en\ repr.\ dom.}$ (5)	^{WRPW,LL}Antp (6)	^{WRPW,LL}Antp (7)	Synth^{YPWM,LL}Antp (8)	Synth^{YPWM,SL}Antp (9)	Synth^{AAAA,LL}Antp (10)	Synth^{AAAA,SL}Antp (11)	Synth^{WRPW,LL}Antp (12)	Synth^{WRPW,SL}Antp (13)
Head involution defects in the larval cuticle	strong	very strong	absent	very strong	medium	medium	medium	medium	strong	very mild	very mild	medium	medium
Repression of *Scr* in the embryo	complete (T1+T2)	complete (T1+T2)	absent	complete (T1+T2)	complete (T1+T2)	complete (T1+T2)	complete (T1+T2)	partial (T2)	complete (T1+T2)	absent	absent	partial (T2)	partial (T2)
Ectopic activation of *Tsh* in the embryo	extensive	extensive	absent	medium	medium	absent	absent	extensive	medium	absent	absent	absent	absent
Repression of *Salm* in the E/A disc	strong	strong	absent	strong	strong	strong	strong	strong	strong	very weak	very weak	absent, disc over-growth	strong
Repression of *dan* in the E/A disc	strong	strong	absent	strong	strong	strong	strong	strong	strong	very weak	very weak	absent, disc over-growth	partial
Ectopic activation of *grn* in the E/A disc	strong	strong	absent	absent	strong	absent	absent	strong	strong	absent	absent	absent	absent
Antenna to tarsus transfor-mation	very strong	very strong	absent	very mild (aristal reduction)	strong	absent	very mild (aristal reduction)	very strong	strong	very mild (aristal reduction)	very mild (aristal reduction)	very mild (A3 over-growth, aristal reduction)	very mild (A3 over-growth, aristal reduction)
Trans-activation in S2-cells	strong	medium	very strong	very strong	not tested	medium	weak	medium	medium	medium	medium	weak	very weak

Table S1. Comparison of Antp gain-of-function phenotypes in the embryo, antennal disc, adult and *Drosophila* S2 cells.

CONCLUSIONS AND OUTLOOK

In the present study we have examined the function of synthetic Hox transcription factors *in vivo* and compared them to their corresponding full-length proteins using various gain-of-function assays at different stages of fly development.

We have used for this purpose synthetic Antp and Scr peptides, lacking the entire N-terminus of the protein, and found them to be functional and to behave in a highly predictable way by exhibiting gene-specific phenotypes at least as strong as the full-length transcription factors. Our study, while assigning a role to the N-terminal portion of Hox factors in transcriptional potency, rather than specificity, *in vivo*, allows us to speculate that the paradigm of the lack of homeotic function of the N-terminus might extend to other Hox transcription factors. This notion would suggest that Hox specificity in general resides only in a small C-terminal fraction of Hox homeoproteins, containing the HD and the YPWM motif.

We have additionally aimed at gaining insight into the molecular mechanisms of Hox mediated transcriptional regulation and the interactions that take place between the HD and nuclear DNA. In order to assure that our measurements are taken under physiological conditions, we induced the synthetic transcription factors in tissue where *Scr* is normally expressed during development and used non-invasive methods with ultimate, single-molecule sensitivity: quantitative imaging and FCS. Our results provided the first *in vivo* binding constant for HD-DNA interactions and developed an easy-to-handle experimental platform for the dissection of molecular numbers and dynamics in live cells.

APPLICATIONS OF SYNTHETIC PEPTIDE TECHNOLOGY FOR USEFUL PURPOSES

It is tempting to assume that our principal findings might apply to synthetic peptides derived from other proteins, which would again be intended to simulate the behavior of their native counterparts (be them transcription factors or not). If this is the case, the development of the knowledge and technology to functionally simulate naturally occurring proteins using synthetic peptides would in turn have important implications for their utilization for useful purposes, including, but not limited to, biomedical applications.

Although the functionality of any less-than-perfect biomolecule (especially when intended to act as a therapeutic agent) requires *in vivo* testing to assure that it

mimics the function of the naturally encountered component, we believe that research and technology pursuing such possibilities are worth applying, since certain advantages usually accompany the use of synthetic compounds, as for instance – when talking about synthetic peptides – smaller size and therefore easier-to-synthesize procedures, easier-to-predict and/or alter molecular properties, but most importantly much higher expected penetration efficiency into target cells or tissue.

Hox proteins are exceptionally important tools for cell penetrating studies. The HD of Antp has been shown to possess cell-penetration properties residing in the third helix (recognition helix, also termed penetratin) of the HD (Derossi et al., 1998; Derossi et al., 1994) (Fig. 8). Other HD-containing proteins, such as En have been shown to be able to travel between cells using their protein domains (Joliot et al., 1998; Prochiantz, 2000). Since the HD and the penetratin sequence are highly conserved, it is thinkable that other HD proteins might display the same properties (Prochiantz, 2000) and this might mean that this mode of signaling has been present in primitive organisms before more elaborate signaling systems have been invented in evolution (Prochiantz, 2008). These findings could potentially be taken advantage of for therapeutic purposes, as in the paradigm of the amplification of CD34 positive hematopoietic precursors by means of HoxB4 internalization (Amsellem et al., 2003), or the inhibition of cell death of dopaminergic neurons using extracellular En in a mouse model for Parkinson's disease (Sonnier et al., 2007).

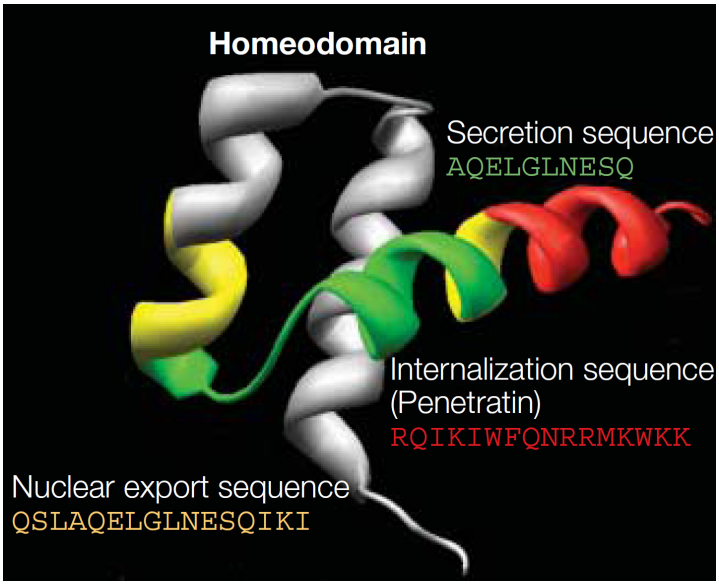

Fig. 7. Functional domains for homeoprotein intercellular transfer. Within the homeodomain, three domains that are required for secretion, internalization and nuclear export have been characterized by loss-of-function (deletion) or gain-of-function (synthetic peptides) studies. The secretion sequence (in green) is part of the nuclear export sequence (in yellow) and its deletion blocks nuclear export. The internalization sequence (in red) has been used as a vector to introduce cargoes into live cells and is therefore also known as Penetratin. It is important to note that the secretion and internalization sequences are distinct and that their presence within the homeodomain makes it impossible to define mutations that would block intercellular transfer without modifying transcription properties (Prochiantz and Joliot, 2003).

Also, the assembly of domains (with known function) from different proteins into novel molecules might present a powerful tool, not yet exploited extensively, to *de novo* engineer new protein functions with remarkable molecular precision. This is an important aim of synthetic biology, which promises to build anew not only proteins, but also signaling networks, or even cells. Ultimate goal of synthetic biologists is the *a priori* prediction of the molecular behavior of a system, upon introduction of synthetic components, or the engineering of a desired function which does not occur normally in nature, for useful purposes (Chen, 2007).

Numerous contributions have been made in the design of peptides for medical applications. A few examples come from the field of antimicrobial peptide development, where the design of synthetic peptides is suggested to be a solution to the poor bioavailability and the high production costs of naturally occurring antimicrobial peptides – reviewed in (Rotem and Mor, 2009) and references therein. The use of peptides as therapeutics has also been demonstrated in vaccine development (Purcell et al., 2007), angiogenesis-related diseases (Sulochana and Ge, 2007), autoimmune diseases (Yusuf-Makagiansar et al., 2002), diabetes (Jain and Chawrai, 2005) and neuroprotection (Gozes and Divinski, 2007).

Among the parallel strategies used is “peptidomimetics”, a practice that aims at mimicking the function of natural proteins, using artificial molecules that resemble them chemically and thus functionally (Goodman et al., 2007; Wells and McClendon, 2007). A lot of effort is being put, for example, in the development of compounds (peptidic or non-peptidic) which bind protein interfaces that participate in protein-protein interactions. The aim in these cases is to disrupt such interactions in a therapeutically relevant way (Wells and McClendon, 2007) and computational tools for modeling such interactions are being developed (Rubinstein and Niv, 2009).

Importantly, several chemical strategies exist nowadays to post-translationally modify synthetic proteins, so that they present as close chemical properties as their functional cellular counterparts (Gamblin et al., 2008), or – in cases where this is tedious to do – even to totally synthesize proteins chemically (Kent, 2009). These chemical strategies enrich the repertoire of molecules used for biomedical applications and provide a solution to the problem of the lack of post-translational modifications in bacterially synthesized peptides. One important application of this practice has been the use of post-translationally modified synthetic peptides (acetylated, lipidated, citrullinated, glycosylated etc.) as biomarkers for the detection

and diagnosis of autoimmune diseases – reviewed in (Papini, 2009). The ability to perform such modifications *in vitro* is essential, since there is growing evidence that autoantibodies recognize post-translationally modified residues on autoantigens (Doyle and Mamula, 2001).

In addition, great progress has been made in the last few years in the design of artificial gene regulatory networks. On the other hand, the engineering of protein regulatory networks has experienced less but considerable progress, most probably due to the complex nature of protein-protein interactions. However, several examples of novel functions using engineered proteins have been reported (see below). In parallel, the generation of synthetic proteins with desired functions using protein domain recombination has important evolutionary implications, since the process of reshuffling of conserved domains has been proposed to be one of the major mechanisms in protein evolution (Lander et al., 2001; Lim, 2002; Pawson and Nash, 2000; Pawson and Nash, 2003).

One example has been the reprogramming of the actin regulatory protein Neuronal Wiskott-Aldrich Syndrome Protein (N-WASP) by changing its autoinhibitory domains to other domains that could respond to the desirable input (Fig. 9), thus constructing an artificial protein which participates in interactions other (but predictable) than its native interactions (Dueber et al., 2003).

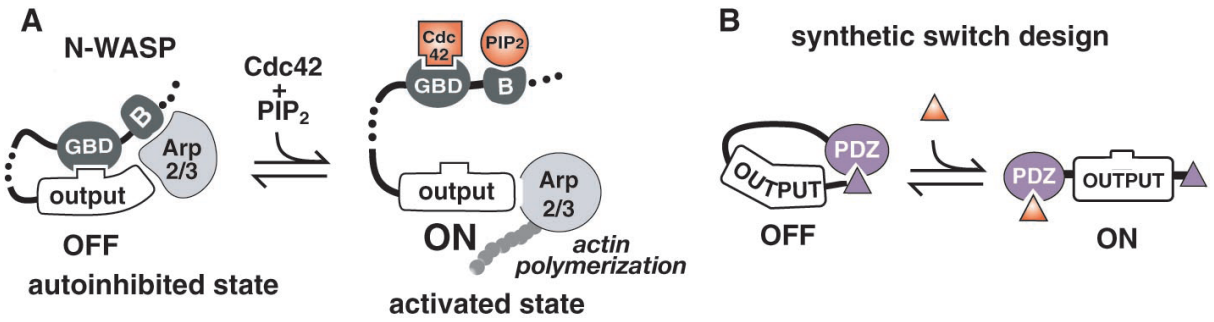

Fig. 9. Design of synthetic switch gated by heterologous ligand. (A) N-WASP is a modular allosteric switch: its output domain constitutively stimulates Arp2/3-mediated actin polymerization but is repressed by autoinhibitory interactions involving two domains, the GTPase-binding domain (GBD) and a basic (B) motif. Input ligands activate by disrupting autoinhibitory interactions: GTP-loaded Cdc42 binds GBD; PIP2 binds B motif. These two inputs act synergistically (*9*, *11*), thus, N-WASP resembles an AND gate. (B) Design strategy for a synthetic single-input switch using N-WASP's output domain and a PDZ domain-ligand pair as heterologous autoinhibitory module (α-syntrophin PDZ; ligand comprising NH2-GVKESLV-COOH; K_d = 8 μM) (Dueber et al., 2003).

The same research group has succeeded also in redesigning synthetic Rho Guanine nucleotide Exchange Factors (GEFs) of the Diffuse B-cell lymphoma (Dbl) family, which normally contain an autoinhibitory domain that represses the function

of their guanine nucleotide exchange domain. Redesign of several GEFs by integrating artificial domains that respond to non-native inputs led to the precise control of physiological responses, as the formation of filopodia or lamellipodia by means of actin cytoskeletal rearrangements (Yeh et al., 2007). Such artificial domains are for example phosphorylation domains which can control the autoinhibitory function of the protein upon phosphorylation. This result suggests that cellular behavior and the spatio-temporal control of cell morphology can – at least to some extent, so far – be reprogrammed and manipulated (Yeh et al., 2007). Similar studies have been performed with the yeast Mitogen Activated Protein (MAP) kinase Ste5 (Park et al., 2003) or the coupling of RTKs to the apoptotic pathway, whereby mitogenic signals could be diverted to stimulate cell death, a result which might have certain anti-oncogenic applications (Howard et al., 2003).

On the other hand, studies demonstrating the non-predictability of synthetic gene or peptide behavior are informative in that they demonstrate that unexpected behavior of such components inside a living cell might also arise, as in the case of a *de novo* engineered bacterial ATP-binding protein. Expression of this peptide in bacterial cells resulted in rapid disruption of the energetic balance and the inhibition of cell division (Stomel et al., 2009).

Finally, it is worth mentioning that synthetic peptides and/or proteins are expected to find use also in an array of non-biomedical applications as for example the development of cheap biodegradable plastics from wheat gluten and dendrimeric thiols (Woerdeman et al., 2004), silk nanofibers for tissue engineering (Li et al., 2006), or photoactive materials based on bacteriorhodopsins (Ho et al., 2004) – reviewed in (Sanford and Kumar, 2005).

REFERENCES

Abzhanov, A. and Kaufman, T. C. (2000). Crustacean (malacostracan) Hox genes and the evolution of the arthropod trunk. *Development* **127**, 2239-49.

Adams, M. D. Celniker, S. E. Holt, R. A. Evans, C. A. Gocayne, J. D. Amanatides, P. G. Scherer, S. E. Li, P. W. Hoskins, R. A. Galle, R. F. et al. (2000). The genome sequence of Drosophila melanogaster. *Science* **287**, 2185-95.

Affolter, M., Percival-Smith, A., Muller, M., Leupin, W. and Gehring, W. J. (1990). DNA binding properties of the purified Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **87**, 4093-7.

Akam, M. (1989). Hox and HOM: homologous gene clusters in insects and vertebrates. *Cell* **57**, 347-9.

Akam, M. E. (1983). The location of Ultrabithorax transcripts in Drosophila tissue sections. *EMBO J* **2**, 2075-84.

Akin, Z. N. and Nazarali, A. J. (2005). Hox genes and their candidate downstream targets in the developing central nervous system. *Cell Mol Neurobiol* **25**, 697-741.

Amsellem, S., Pflumio, F., Bardinet, D., Izac, B., Charneau, P., Romeo, P. H., Dubart-Kupperschmitt, A. and Fichelson, S. (2003). Ex vivo expansion of human hematopoietic stem cells by direct delivery of the HOXB4 homeoprotein. *Nat Med* **9**, 1423-7.

Andrew, D. J., Horner, M. A., Petitt, M. G., Smolik, S. M. and Scott, M. P. (1994). Setting limits on homeotic gene function: restraint of Sex combs reduced activity by teashirt and other homeotic genes. *EMBO J* **13**, 1132-44.

Arias, A. M. (2008). Drosophila melanogaster and the development of biology in the 20th century. *Methods Mol Biol* **420**, 1-25.

Bacia, K. and Schwille, P. (2003). A dynamic view of cellular processes by in vivo fluorescence auto- and cross-correlation spectroscopy. *Methods* **29**, 74-85.

Bale, A. E. (2002). Hedgehog signaling and human disease. *Annu Rev Genomics Hum Genet* **3**, 47-65.

Barth, A. I. and Nelson, W. J. (2002). What can humans learn from flies about adenomatous polyposis coli? *Bioessays* **24**, 771-4.

Beeman, R. W., Stuart, J. J., Haas, M. S. and Denell, R. E. (1989). Genetic analysis of the homeotic gene complex (HOM-C) in the beetle Tribolium castaneum. *Dev Biol* **133**, 196-209.

Bellen, H. J. (1998). The fruit fly: a model organism to study the genetics of alcohol abuse and addiction? *Cell* **93**, 909-12.

Bellen, H. J., O'Kane, C. J., Wilson, C., Grossniklaus, U., Pearson, R. K. and Gehring, W. J. (1989). P-element-mediated enhancer detection: a versatile method to study development in Drosophila. *Genes Dev* **3**, 1288-300.

Bellen, H. J., Wilson, C., Gibson, G., Grossniklaus, U., Pearson, R. K., O'Kane, C. and Gehring, W. J. (1990). P-element-mediated enhancer detection allows rapid identification of developmentally regulated genes and cell specific markers in Drosophila. *J Physiol (Paris)* **84**, 33-41.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., Lewis, E. B. and Hogness, D. S. (1983a). Molecular Genetics of the Bithorax Complex in Drosophila melanogaster. *Science* **221**, 23-29.

Bender, W., Spierer, P. and Hogness, D. S. (1983b). Chromosomal walking and jumping to isolate DNA from the Ace and rosy loci and the bithorax complex in Drosophila melanogaster. *J Mol Biol* **168**, 17-33.

Berger, M. F., Badis, G., Gehrke, A. R., Talukder, S., Philippakis, A. A., Pena-Castillo, L., Alleyne, T. M., Mnaimneh, S., Botvinnik, O. B., Chan, E. T. et al. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**, 1266-76.

Berry, M. and Gehring, W. (2000). Phosphorylation status of the SCR homeodomain determines its functional activity: essential role for protein phosphatase 2A,B'. *EMBO J* **19**, 2946-57.

Berthelsen, J., Kilstrup-Nielsen, C., Blasi, F., Mavilio, F. and Zappavigna, V. (1999). The subcellular localization of PBX1 and EXD proteins depends on nuclear import and export signals and is modulated by association with PREP1 and HTH. *Genes Dev* **13**, 946-53.

Berthelsen, J., Zappavigna, V., Mavilio, F. and Blasi, F. (1998). Prep1, a novel functional partner of Pbx proteins. *EMBO J* **17**, 1423-33.

Bessa, J., Tavares, M. J., Santos, J., Kikuta, H., Laplante, M., Becker, T. S., Gomez-Skarmeta, J. L. and Casares, F. (2008). meis1 regulates cyclin D1 and c-myc expression, and controls the proliferation of the multipotent cells in the early developing zebrafish eye. *Development* **135**, 799-803.

Bier, E. (2005). Drosophila, the golden bug, emerges as a tool for human genetics. *Nat Rev Genet* **6**, 9-23.

Billeter, M. and Wuthrich, K. (1993). Model studies relating nuclear magnetic resonance data with the three-dimensional structure of protein-DNA complexes. *J Mol Biol* **234**, 1094-7.

Bodai, L., Pallos, J., Thompson, L. M. and Marsh, J. L. (2003). Altered protein acetylation in polyglutamine diseases. *Curr Med Chem* **10**, 2577-87.

Boncinelli, E., Acampora, D., Pannese, M., D'Esposito, M., Somma, R., Gaudino, G., Stornaiuolo, A., Cafiero, M., Faiella, A. and Simeone, A. (1989). Organization of human class I homeobox genes. *Genome* **31**, 745-56.

Bonini, N. M. and Fortini, M. E. (2003). Human neurodegenerative disease modeling using Drosophila. *Annu Rev Neurosci* **26**, 627-56.

Boutros, M. and Perrimon, N. (2000). Drosophila genome takes flight. *Nat Cell Biol* **2**, E53-4.

Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-15.

Brumby, A. M. and Richardson, H. E. (2003). scribble mutants cooperate with oncogenic Ras or Notch to cause neoplastic overgrowth in Drosophila. *EMBO J* **22**, 5769-79.

Burglin, T. R. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res* **25**, 4173-80.

Calleja, M., Moreno, E., Pelaz, S. and Morata, G. (1996). Visualization of gene expression in living adult Drosophila. *Science* **274**, 252-5.

Camacho, A., Korn, K., Damond, M., Cajot, J. F., Litborn, E., Liao, B., Thyberg, P., Winter, H., Honegger, A., Gardellin, P. et al. (2004). Direct quantification of mRNA expression levels using single molecule detection. *J Biotechnol* **107**, 107-14.

Carpenter, E. M. (2002). Hox genes and spinal cord development. *Dev Neurosci* **24**, 24-34.

Carroll, S. B., Weatherbee, S. D. and Langeland, J. A. (1995). Homeotic genes and the regulation and evolution of insect wing number. *Nature* **375**, 58-61.

Casares, F. and Mann, R. S. (1998). Control of antennal versus leg development in Drosophila. *Nature* **392**, 723-6.

Chabbert, M., Hillen, W., Hansen, D., Takahashi, M. and Bousquet, J. A. (1992). Structural analysis of the operator binding domain of Tn10-encoded Tet repressor: a time-resolved fluorescence and anisotropy study. *Biochemistry* **31**, 1951-60.
Chan, S. K. and Mann, R. S. (1993). The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes Dev* **7**, 796-811.
Chan, S. K. and Mann, R. S. (1996). A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc Natl Acad Sci U S A* **93**, 5223-8.
Chang, C. P., Shen, W. F., Rozenfeld, S., Lawrence, H. J., Largman, C. and Cleary, M. L. (1995). Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev* **9**, 663-74.
Chen, C. S. (2007). Bait and switch: synthetic GEFs divert an input signal to diverse cellular responses. *Dev Cell* **13**, 9-10.
Chesebro, J., Hrycaj, S., Mahfooz, N. and Popadic, A. (2009). Diverging functions of Scr between embryonic and post-embryonic development in a hemimetabolous insect, Oncopeltus fasciatus. *Dev Biol* **329**, 142-51.
Christ, B., Schmidt, C., Huang, R., Wilting, J. and Brand-Saberi, B. (1998). Segmentation of the vertebrate body. *Anat Embryol (Berl)* **197**, 1-8.
Cook, C. E., Jimenez, E., Akam, M. and Salo, E. (2004). The Hox gene complement of acoel flatworms, a basal bilaterian clade. *Evol Dev* **6**, 154-63.
Cook, J., Holtom, G. and Lu, P. (1990). Detection of protein-DNA complex formation by time-resolved fluorescence depolarization of bound ethidium bromide. *Anal Biochem* **190**, 331-9.
Czerny, T., Halder, G., Kloter, U., Souabni, A., Gehring, W. J. and Busslinger, M. (1999). twin of eyeless, a second Pax-6 gene of Drosophila, acts upstream of eyeless in the control of eye development. *Mol Cell* **3**, 297-307.
Dasen, J. S., De Camilli, A., Wang, B., Tucker, P. W. and Jessell, T. M. (2008). Hox repertoires for motor neuron diversity and connectivity gated by a single accessory factor, FoxP1. *Cell* **134**, 304-16.
De Celis, J. F. (2003). Pattern formation in the Drosophila wing: The development of the veins. *Bioessays* **25**, 443-51.
Derossi, D., Chassaing, G. and Prochiantz, A. (1998). Trojan peptides: the penetratin system for intracellular delivery. *Trends Cell Biol* **8**, 84-7.
Derossi, D., Joliot, A. H., Chassaing, G. and Prochiantz, A. (1994). The third helix of the Antennapedia homeodomain translocates through biological membranes. *J Biol Chem* **269**, 10444-50.
Doucleff, M. and Clore, G. M. (2008). Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1. *Proc Natl Acad Sci U S A* **105**, 13871-6.
Dow, L. E., Brumby, A. M., Muratore, R., Coombe, M. L., Sedelies, K. A., Trapani, J. A., Russell, S. M., Richardson, H. E. and Humbert, P. O. (2003). hScrib is a functional homologue of the Drosophila tumour suppressor Scribble. *Oncogene* **22**, 9225-30.
Doyle, H. A. and Mamula, M. J. (2001). Post-translational protein modifications in antigen recognition and autoimmunity. *Trends Immunol* **22**, 443-9.
Driscoll, M. and Gerstbrein, B. (2003). Dying for a cause: invertebrate genetics takes on human neurodegeneration. *Nat Rev Genet* **4**, 181-94.

Duboule, D. and Dolle, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J* **8**, 1497-505.
Duboule, D. and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet* **10**, 358-64.
Dueber, J. E., Yeh, B. J., Chak, K. and Lim, W. A. (2003). Reprogramming control of an allosteric signaling switch through modular recombination. *Science* **301**, 1904-8.
Ehrenberg, M. and Rigler, R. (1974). Rotational Brownian-Motion and Fluorescence Intensity Fluctuations. *Chemical Physics* **4**, 390-401.
Eigen, M. and Rigler, R. (1994). Sorting Single Molecules - Application to Diagnostics and Evolutionary Biotechnology. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5740-5747.
Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E. and Beachy, P. A. (1994). The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins. *EMBO J* **13**, 3551-60.
Elson, E. L. and Magde, D. (1974). Fluorescence Correlation Spectroscopy .1. Conceptual Basis and Theory. *Biopolymers* **13**, 1-27.
Emerald, B. S., Curtiss, J., Mlodzik, M. and Cohen, S. M. (2003). Distal antenna and distal antenna related encode nuclear proteins containing pipsqueak motifs involved in antenna development in Drosophila. *Development* **130**, 1171-80.
Feany, M. B. and Bender, W. W. (2000). A Drosophila model of Parkinson's disease. *Nature* **404**, 394-8.
Finnerty, J. R. (2005). Did internal transport, rather than directed locomotion, favor the evolution of bilateral symmetry in animals? *Bioessays* **27**, 1174-80.
Finnerty, J. R. and Martindale, M. Q. (1997). Homeoboxes in sea anemones (Cnidaria:Anthozoa): a PCR-based survey of Nematostella vectensis and Metridium senile. *Biol Bull* **193**, 62-76.
Finnerty, J. R. and Martindale, M. Q. (1999). Ancient origins of axial patterning genes: Hox genes and ParaHox genes in the Cnidaria. *Evol Dev* **1**, 16-23.
Finnerty, J. R., Pang, K., Burton, P., Paulson, D. and Martindale, M. Q. (2004). Origins of bilateral symmetry: Hox and dpp expression in a sea anemone. *Science* **304**, 1335-7.
Fisher, A. L., Ohsako, S. and Caudy, M. (1996). The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. *Mol Cell Biol* **16**, 2670-7.
Fradkov, A. F., Chen, Y., Ding, L., Barsova, E. V., Matz, M. V. and Lukyanov, S. A. (2000). Novel fluorescent protein from Discosoma coral and its mutants possesses a unique far-red fluorescence. *FEBS Lett* **479**, 127-30.
Friedrich, M. and Tautz, D. (1995). Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. *Nature* **376**, 165-7.
Furukubo-Tokunaga, K., Flister, S. and Gehring, W. J. (1993). Functional specificity of the Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **90**, 6360-4.
Furukubo-Tokunaga, K., Muller, M., Affolter, M., Pick, L., Kloter, U. and Gehring, W. J. (1992). In vivo analysis of the helix-turn-helix motif of the fushi tarazu homeo domain of Drosophila melanogaster. *Genes Dev* **6**, 1082-96.
Galant, R., Walsh, C. M. and Carroll, S. B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* **129**, 3115-26.

Gamblin, D. P., van Kasteren, S. I., Chalker, J. M. and Davis, B. G. (2008). Chemical approaches to mapping the function of post-translational modifications. *FEBS J* **275**, 1949-59.
Gangloff, Y. G., Pointud, J. C., Thuault, S., Carre, L., Romier, C., Muratoglu, S., Brand, M., Tora, L., Couderc, J. L. and Davidson, I. (2001). The TFIID components human TAF(II)140 and Drosophila BIP2 (TAF(II)155) are novel metazoan homologues of yeast TAF(II)47 containing a histone fold and a PHD finger. *Mol Cell Biol* **21**, 5109-21.
Garber, R. L., Kuroiwa, A. and Gehring, W. J. (1983). Genomic and cDNA clones of the homeotic locus Antennapedia in Drosophila. *EMBO J* **2**, 2027-36.
Garcia-Bellido, A. and Merriam, J. R. (1971). Parameters of the wing imaginal disc development of Drosophila melanogaster. *Dev Biol* **24**, 61-87.
Gauchat, D., Mazet, F., Berney, C., Schummer, M., Kreger, S., Pawlowski, J. and Galliot, B. (2000). Evolution of Antp-class genes and differential expression of Hydra Hox/paraHox genes in anterior patterning. *Proc Natl Acad Sci U S A* **97**, 4493-8.
Gebelein, B., McKay, D. J. and Mann, R. S. (2004). Direct integration of Hox and segmentation gene inputs during Drosophila development. *Nature* **431**, 653-9.
Gehring, W. J. (1987). Homeo boxes in the study of development. *Science* **236**, 1245-52.
Gehring, W. J. (1998). Master control genes in development and evolution: the homeonox story. New Haven, CT: Yale University Press.
Gehring, W. J., Kloter, U. and Suga, H. (2009). Evolution of the Hox gene complex from an evolutionary ground state. *Curr Top Dev Biol* **88**, 35-61.
Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994). Homeodomain-DNA recognition. *Cell* **78**, 211-23.
Gibson, G., Schier, A., LeMotte, P. and Gehring, W. J. (1990). The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**, 1087-103.
Giribet, G., Edgecombe, G. D. and Wheeler, W. C. (2001). Arthropod phylogeny based on eight molecular loci and morphology. *Nature* **413**, 157-61.
Goldberg, M. L., Sheen, J. Y., Gehring, W. J. and Green, M. M. (1983). Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the Drosophila melanogaster genome. *Proc Natl Acad Sci U S A* **80**, 5017-5021.
Golic, K. G. and Lindquist, S. (1989). The FLP recombinase of yeast catalyzes site-specific recombination in the Drosophila genome. *Cell* **59**, 499-509.
Goodman, C. M., Choi, S., Shandler, S. and DeGrado, W. F. (2007). Foldamers as versatile frameworks for the design and evolution of function. *Nat Chem Biol* **3**, 252-62.
Gozes, I. and Divinski, I. (2007). NAP, a neuroprotective drug candidate in clinical trials, stimulates microtubule assembly in the living cell. *Curr Alzheimer Res* **4**, 507-9.
Graham, A., Papalopulu, N., Lorimer, J., McVey, J. H., Tuddenham, E. G. and Krumlauf, R. (1988). Characterization of a murine homeo box gene, Hox-2.6, related to the Drosophila Deformed gene. *Genes Dev* **2**, 1424-38.
Grieder, N. C., Morata, G., Affolter, M. and Gehring, W. J. (2009). Spalt major controls the development of the notum and of wing hinge primordia of the Drosophila melanogaster wing imaginal disc. *Dev Biol* **329**, 315-26.
Grienenberger, A., Merabet, S., Manak, J., Iltis, I., Fabre, A., Berenger, H., Scott, M. P., Pradel, J. and Graba, Y. (2003). Tgfbeta signaling acts on a Hox response

element to confer specificity and diversity to Hox protein function. *Development* **130**, 5445-55.
Grossniklaus, U., Bellen, H. J., Wilson, C. and Gehring, W. J. (1989). P-element-mediated enhancer detection applied to the study of oogenesis in Drosophila. *Development* **107**, 189-200.
Hafen, E., Levine, M., Garber, R. L. and Gehring, W. J. (1983). An improved in situ hybridization method for the detection of cellular RNAs in Drosophila tissue sections and its application for localizing transcripts of the homeotic Antennapedia gene complex. *EMBO J* **2**, 617-623.
Harvey, K. F., Pfleger, C. M. and Hariharan, I. K. (2003). The Drosophila Mst ortholog, hippo, restricts growth and cell proliferation and promotes apoptosis. *Cell* **114**, 457-67.
Haustein, E. and Schwille, P. (2007). Fluorescence correlation spectroscopy: novel variations of an established technique. *Annu Rev Biophys Biomol Struct* **36**, 151-69.
Haywood, A. F. and Staveley, B. E. (2004). Parkin counteracts symptoms in a Drosophila model of Parkinson's disease. *BMC Neurosci* **5**, 14.
Ho, D., Chu, B., Lee, H. and Montemagno, C. D. (2004). Protein-driven energy transduction across polymeric biomembranes. *Nanotechnology* **15**, 1084-1094.
Hoey, T. and Levine, M. (1988). Divergent homeo box proteins recognize similar DNA sequences in Drosophila. *Nature* **332**, 858-61.
Holland, P. W., Garcia-Fernandez, J., Williams, N. A. and Sidow, A. (1994). Gene duplications and the origins of vertebrate development. *Dev Suppl*, 125-33.
Hoskins, R. A., Carlson, J. W., Kennedy, C., Acevedo, D., Evans-Holm, M., Frise, E., Wan, K. H., Park, S., Mendez-Lago, M., Rossi, F. et al. (2007). Sequence finishing and mapping of Drosophila melanogaster heterochromatin. *Science* **316**, 1625-8.
Howard, K. and Ingham, P. (1986). Regulatory interactions between the segmentation genes fushi tarazu, hairy, and engrailed in the Drosophila blastoderm. *Cell* **44**, 949-57.
Howard, P. L., Chia, M. C., Del Rizzo, S., Liu, F. F. and Pawson, T. (2003). Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. *Proc Natl Acad Sci U S A* **100**, 11267-72.
Hrycaj, S., Chesebro, J. and Popadic, A. (2010). Functional analysis of Scr during embryonic and post-embryonic development in the cockroach, Periplaneta americana. *Dev Biol*.
Hueber, S. D. and Lohmann, I. (2008). Shaping segments: Hox gene function in the genomic age. *Bioessays* **30**, 965-79.
Huet, A., Parlakian, A., Arnaud, M. C., Glandieres, J. M., Valat, P., Fermandjian, S., Paulin, D., Alpert, B. and Zentz, C. (2005). Mechanism of binding of serum response factor to serum response element. *FEBS J* **272**, 3105-19.
Huson, D. H. and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**, 254-67.
Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in Drosophila. *Nature* **335**, 25-34.
Ishizuka, A., Siomi, M. C. and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* **16**, 2497-508.
Jain, R. and Chawrai, S. (2005). Advancements in the anti-diabetes chemotherapeutics based on amino acids, peptides, and peptidomimetics. *Mini Rev Med Chem* **5**, 469-77.

Janody, F., Sturny, R., Schaeffer, V., Azou, Y. and Dostatni, N. (2001). Two distinct domains of Bicoid mediate its transcriptional downregulation by the Torso pathway. *Development* **128**, 2281-90.
Jiang, Y., Shi, H., Amin, N. M., Sultan, I. and Liu, J. (2008). Mesodermal expression of the C. elegans HMX homolog mls-2 requires the PBC homolog CEH-20. *Mech Dev* **125**, 451-61.
Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K. and Warren, S. T. (2003). RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in Drosophila. *Neuron* **39**, 739-47.
Johnson, F. B., Parker, E. and Krasnow, M. A. (1995). Extradenticle protein is a selective cofactor for the Drosophila homeotics: role of the homeodomain and YPWM amino acid motif in the interaction. *Proc Natl Acad Sci U S A* **92**, 739-43.
Johnston, R. N., Pai, S. B. and Pai, R. B. (1992). The origin of the cancer cell: oncogeny reverses phylogeny. *Biochem Cell Biol* **70**, 831-4.
Joliot, A., Maizel, A., Rosenberg, D., Trembleau, A., Dupas, S., Volovitch, M. and Prochiantz, A. (1998). Identification of a signal sequence necessary for the unconventional secretion of Engrailed homeoprotein. *Curr Biol* **8**, 856-63.
Joshi, R., Passner, J. M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A., Jacob, V., Aggarwal, A. K., Honig, B. and Mann, R. S. (2007). Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* **131**, 530-43.
Justice, R. W., Zilian, O., Woods, D. F., Noll, M. and Bryant, P. J. (1995). The Drosophila tumor suppressor gene warts encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation. *Genes Dev* **9**, 534-46.
Kaufman, T. C., Lewis, R. and Wakimoto, B. (1980). Cytogenetic Analysis of Chromosome 3 in DROSOPHILA MELANOGASTER: The Homoeotic Gene Complex in Polytene Chromosome Interval 84a-B. *Genetics* **94**, 115-133.
Kent, S. B. (2009). Total chemical synthesis of proteins. *Chem Soc Rev* **38**, 338-51.
Khrapunov, S., Pastor, N. and Brenowitz, M. (2002). Solution structural studies of the Saccharomyces cerevisiae TATA binding protein (TBP). *Biochemistry* **41**, 9559-71.
King, T. J. and Briggs, R. (1956). Serial transplantation of embryonic nuclei. *Cold Spring Harb Symp Quant Biol* **21**, 271-90.
Kinjo, M. and Rigler, R. (1995). Ultrasensitive hybridization analysis using fluorescence correlation spectroscopy. *Nucleic Acids Res* **23**, 1795-9.
Kissinger, C. R., Liu, B. S., Martin-Blanco, E., Kornberg, T. B. and Pabo, C. O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. *Cell* **63**, 579-90.
Kmita-Cunisse, M., Loosli, F., Bierne, J. and Gehring, W. J. (1998). Homeobox genes in the ribbonworm Lineus sanguineus: evolutionary implications. *Proc Natl Acad Sci U S A* **95**, 3030-5.
Kmita, G., Baranska, M., Ciaston-Przeclawska, E., Ceranska-Goszczynska, H. and Urmanska, W. (1998). [Psychological aspects of high risk pregnancy in hospitalized women -- toward the change of the existing model of care]. *Ginekol Pol* **69**, 1025-31.
Kmita, M. and Duboule, D. (2003). Organizing axes in time and space; 25 years of colinear tinkering. *Science* **301**, 331-3.

Kobayashi, T., Okamoto, N., Sawasaki, T. and Endo, Y. (2004). Detection of protein-DNA interactions in crude cellular extracts by fluorescence correlation spectroscopy. *Anal Biochem* **332**, 58-66.
Kohlhase, J. (2000). SALL1 mutations in Townes-Brocks syndrome and related disorders. *Hum Mutat* **16**, 460-6.
Koppel, D. E. (1974). Statistical accuracy in fluorescence correlation spectroscopy. *Phys. Rev. A* **10**, 1938-1945.
Koppel, D. E., Axelrod, D., Schlessinger, J., Elson, E. L. and Webb, W. W. (1976). Dynamics of Fluorescence Marker Concentration as a Probe of Mobility. *Biophysical Journal* **16**, 1315-1329.
Korn, K., Gardellin, P., Liao, B., Amacker, M., Bergstrom, A., Bjorkman, H., Camacho, A., Dorhofer, S., Dorre, K., Enstrom, J. et al. (2003). Gene expression analysis using single molecule detection. *Nucleic Acids Res* **31**, e89.
Kosman, D., Mizutani, C. M., Lemons, D., Cox, W. G., McGinnis, W. and Bier, E. (2004). Multiplex detection of RNA expression in Drosophila embryos. *Science* **305**, 846.
Krumlauf, R., Chapman, V. M., Hammer, R. E., Brinster, R. and Tilghman, S. M. (1986). Differential expression of alpha-fetoprotein genes on the inactive X chromosome in extraembryonic and somatic tissues of a transgenic mouse line. *Nature* **319**, 224-6.
Kuhnlein, R. P., Frommer, G., Friedrich, M., Gonzalez-Gaitan, M., Weber, A., Wagner-Bernholz, J. F., Gehring, W. J., Jackle, H. and Schuh, R. (1994). spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. *EMBO J* **13**, 168-79.
Kumar, P. and Nazarali, A. J. (2001). Characterization of Hoxd1 protein-DNA-binding specificity using affinity chromatography and random DNA oligomer selection. *Cell Mol Neurobiol* **21**, 369-88.
Kurant, E., Pai, C. Y., Sharf, R., Halachmi, N., Sun, Y. H. and Salzberg, A. (1998). Dorsotonals/homothorax, the Drosophila homologue of meis1, interacts with extradenticle in patterning of the embryonic PNS. *Development* **125**, 1037-48.
Lander, E. S. Linton, L. M. Birren, B. Nusbaum, C. Zody, M. C. Baldwin, J. Devon, K. Dewar, K. Doyle, M. FitzHugh, W. et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.
LaRonde-LeBlanc, N. A. and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. *Genes Dev* **17**, 2060-72.
Laughon, A. and Scott, M. P. (1984). Sequence of a Drosophila segmentation gene: protein structure homology with DNA-binding proteins. *Nature* **310**, 25-31.
Laurent, A., Bihan, R., Deschamps, S., Guerrier, D., Dupe, V., Omilli, F., Burel, A. and Pellerin, I. (2007). Identification of a new type of PBX1 partner that contains zinc finger motifs and inhibits the binding of HOXA9-PBX1 to DNA. *Mech Dev* **124**, 364-76.
Lawrence, P. A. (1981). A general cell marker for clonal analysis of Drosophila development. *J Embryol Exp Morphol* **64**, 321-32.
LeMotte, P. K., Kuroiwa, A., Fessler, L. I. and Gehring, W. J. (1989). The homeotic gene Sex Combs Reduced of Drosophila: gene structure and embryonic expression. *EMBO J* **8**, 219-27.

Levine, M., Rubin, G. M. and Tjian, R. (1984). Human DNA sequences homologous to a protein coding region conserved between homeotic genes of Drosophila. *Cell* **38**, 667-73.
Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**, 565-70.
Lewis, E. B. (1994). Homeosis: the first 100 years. *Trends Genet* **10**, 341-3.
Li, C., Vepari, C., Jin, H. J., Kim, H. J. and Kaplan, D. L. (2006). Electrospun silk-BMP-2 scaffolds for bone tissue engineering. *Biomaterials* **27**, 3115-24.
Lim, W. A. (2002). The modular logic of signaling proteins: building allosteric switches from simple binding domains. *Curr Opin Struct Biol* **12**, 61-8.
Lin, L. and McGinnis, W. (1992). Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev* **6**, 1071-81.
Lipshitz, H. D. (2004). From fruit flies to fallout: Ed Lewis and his science. *J Genet* **83**, 201-18.
Liu, Y., Matthews, K. S. and Bondos, S. E. (2009). Internal regulatory interactions determine DNA binding specificity by a Hox transcription factor. *J Mol Biol* **390**, 760-74.
Lohr, U. and Pick, L. (2005). Cofactor-interaction motifs and the cooption of a homeotic Hox protein into the segmentation pathway of Drosophila melanogaster. *Curr Biol* **15**, 643-9.
Lu, Q. and Kamps, M. P. (1996). Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: proposal for a model of a Pbx1-Hox-DNA complex. *Mol Cell Biol* **16**, 1632-40.
Magde, D., Webb, W. W. and Elson, E. (1972). Thermodynamic Fluctuations in a Reacting System - Measurement by Fluorescence Correlation Spectroscopy. *Physical Review Letters* **29**, 705-&.
Mann, R. S. and Affolter, M. (1998). Hox proteins meet more partners. *Curr Opin Genet Dev* **8**, 423-9.
Mann, R. S. and Chan, S. K. (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. *Trends Genet* **12**, 258-62.
Mann, R. S. and Hogness, D. S. (1990). Functional dissection of Ultrabithorax proteins in D. melanogaster. *Cell* **60**, 597-610.
Mann, R. S., Lelli, K. M. and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. *Curr Top Dev Biol* **88**, 63-101.
Mann, R. S. and Morata, G. (2000). The developmental and molecular biology of genes that subdivide the body of Drosophila. *Annu Rev Cell Dev Biol* **16**, 243-71.
Manuel, M., Jager, M., Murienne, J., Clabaut, C. and Le Guyader, H. (2006). Hox genes in sea spiders (Pycnogonida) and the homology of arthropod head segments. *Dev Genes Evol* **216**, 481-91.
Margulies, C., Tully, T. and Dubnau, J. (2005). Deconstructing memory in Drosophila. *Curr Biol* **15**, R700-13.
Maroteaux, L., Campanelli, J. T. and Scheller, R. H. (1988). Synuclein: a neuron-specific protein localized to the nucleus and presynaptic nerve terminal. *J Neurosci* **8**, 2804-15.
Matthews, K. A., Kaufman, T. C. and Gelbart, W. M. (2005). Research resources for Drosophila: the expanding universe. *Nat Rev Genet* **6**, 179-93.
Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L. and Lukyanov, S. A. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotechnol* **17**, 969-73.

Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Zappavigna, V., Acampora, D., Poiana, G., Russo, G., Peschle, C. and Boncinelli, E. (1986). Differential and stage-related expression in embryonic tissues of a new human homoeobox gene. *Nature* **324**, 664-8.
McCormick, A., Core, N., Kerridge, S. and Scott, M. P. (1995). Homeotic response elements are tightly linked to tissue-specific elements in a transcriptional enhancer of the teashirt gene. *Development* **121**, 2799-812.
McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. and Gehring, W. J. (1984a). A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. *Cell* **37**, 403-8.
McGinnis, W., Hart, C. P., Gehring, W. J. and Ruddle, F. H. (1984b). Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of Drosophila. *Cell* **38**, 675-80.
McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.
McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. and Gehring, W. J. (1984c). A conserved DNA sequence in homoeotic genes of the Drosophila Antennapedia and bithorax complexes. *Nature* **308**, 428-33.
Merabet, S., Kambris, Z., Capovilla, M., Berenger, H., Pradel, J. and Graba, Y. (2003). The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Dev Cell* **4**, 761-8.
Michelman-Ribeiro, A., Mazza, D., Rosales, T., Stasevich, T. J., Boukari, H., Rishi, V., Vinson, C., Knutson, J. R. and McNally, J. G. (2009). Direct measurement of association and dissociation rates of DNA binding in live cells by fluorescence correlation spectroscopy. *Biophys J* **97**, 337-46.
Miyawaki, A., Sawano, A. and Kogure, T. (2003). Lighting up cells: labelling proteins with fluorophores. *Nat Cell Biol* **Suppl**, S1-7.
Moens, C. B. and Selleri, L. (2006). Hox cofactors in vertebrate development. *Dev Biol* **291**, 193-206.
Moerner, W. E. and Kador, L. (1989). Optical detection and spectroscopy of single molecules in a solid. *Phys Rev Lett* **62**, 2535-2538.
Morgan, T. H. (1910). Sex Limited Inheritance in Drosophila. *Science* **32**, 120-122.
Muller, M., Affolter, M., Leupin, W., Otting, G., Wuthrich, K. and Gehring, W. J. (1988). Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J* **7**, 4299-304.
Mulligan, L. M. (2001). Multiple endocrine neoplasia type 2: molecular aspects. *Front Horm Res* **28**, 81-102.
Muqit, M. M. and Feany, M. B. (2002). Modelling neurodegenerative diseases in Drosophila: a fruitful approach? *Nat Rev Neurosci* **3**, 237-43.
Muto, H., Nagao, I., Demura, T., Fukuda, H., Kinjo, M. and Yamamoto, K. T. (2006). Fluorescence cross-correlation analyses of the molecular interaction between an Aux/IAA protein, MSG2/IAA19, and protein-protein interaction domains of auxin response factors of arabidopsis expressed in HeLa cells. *Plant Cell Physiol* **47**, 1095-101.
Myers, E. W., Sutton, G. G., Delcher, A. L., Dew, I. M., Fasulo, D. P., Flanigan, M. J., Kravitz, S. A., Mobarry, C. M., Reinert, K. H., Remington, K. A. et al. (2000). A whole-genome assembly of Drosophila. *Science* **287**, 2196-204.
Nandi, C. K., Parui, P. P., Brutschy, B., Scheffer, U. and Gobel, M. (2008). Fluorescence correlation spectroscopy at single molecule level on the Tat-TAR complex and its inhibitors. *Biopolymers* **89**, 17-25.

Neuteboom, S. T., Peltenburg, L. T., van Dijk, M. A. and Murre, C. (1995). The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins. *Proc Natl Acad Sci U S A* **92**, 9166-70.
Noyes, M. B., Christensen, R. G., Wakabayashi, A., Stormo, G. D., Brodsky, M. H. and Wolfe, S. A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* **133**, 1277-89.
Nusslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. *Nature* **287**, 795-801.
O'Connor, M. B., Binari, R., Perkins, L. A. and Bender, W. (1988). Alternative RNA products from the Ultrabithorax domain of the bithorax complex. *EMBO J* **7**, 435-45.
O'Kane, C. J. and Gehring, W. J. (1987). Detection in situ of genomic regulatory elements in Drosophila. *Proc Natl Acad Sci U S A* **84**, 9123-7.
Octobre, G., Lemercier, C., Khochbin, S., Robert-Nicoud, M. and Souchier, C. (2005). Monitoring the interaction between DNA and a transcription factor (MEF2A) using fluorescence correlation spectroscopy. *C R Biol* **328**, 1033-40.
Orrit, M. and Bernard, J. (1990). Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal. *Phys Rev Lett* **65**, 2716-2719.
Otting, G., Qian, Y. Q., Muller, M., Affolter, M., Gehring, W. and Wuthrich, K. (1988). Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. *EMBO J* **7**, 4305-9.
Pagliarini, R. A. and Xu, T. (2003). A genetic screen in Drosophila for metastatic behavior. *Science* **302**, 1227-31.
Pan, D., Dong, J., Zhang, Y. and Gao, X. (2004). Tuberous sclerosis complex: from Drosophila to human disease. *Trends Cell Biol* **14**, 78-85.
Papadopoulos, D. K., Vukojevic, V., Adachi, Y., Terenius, L., Rigler, R. and Gehring, W. J. (2010). Function and specificity of synthetic Hox transcription factors in vivo. *Proc Natl Acad Sci U S A* **107**, 4087-92.
Papini, A. M. (2009). The use of post-translationally modified peptides for detection of biomarkers of immune-mediated diseases. *J Pept Sci* **15**, 621-8.
Park, S. H., Zarrinpar, A. and Lim, W. A. (2003). Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**, 1061-4.
Paroush, Z., Finley, R. L., Jr., Kidd, T., Wainwright, S. M., Ingham, P. W., Brent, R. and Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**, 805-15.
Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S. and Aggarwal, A. K. (1999). Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* **397**, 714-9.
Pawson, T. and Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes Dev* **14**, 1027-47.
Pawson, T. and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445-52.
Pearson, J. C., Lemons, D. and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. *Nat Rev Genet* **6**, 893-904.
Peifer, M. and Wieschaus, E. (1990). Mutations in the Drosophila gene extradenticle affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev* **4**, 1209-23.
Perrimon, N. (1998). Creating mosaics in Drosophila. *Int J Dev Biol* **42**, 243-7.

Phelan, M. L., Rambaldi, I. and Featherstone, M. S. (1995). Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol Cell Biol* **15**, 3989-97.
Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L. and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* **96**, 587-97.
Plaza, S., Prince, F., Adachi, Y., Punzo, C., Cribbs, D. L. and Gehring, W. J. (2008). Cross-regulatory protein-protein interactions between Hox and Pax transcription factors. *Proc Natl Acad Sci U S A* **105**, 13439-44.
Plaza, S., Prince, F., Jaeger, J., Kloter, U., Flister, S., Benassayag, C., Cribbs, D. and Gehring, W. J. (2001). Molecular basis for the inhibition of Drosophila eye development by Antennapedia. *EMBO J* **20**, 802-11.
Prince, F., Katsuyama, T., Oshima, Y., Plaza, S., Resendez-Perez, D., Berry, M., Kurata, S. and Gehring, W. J. (2008). The YPWM motif links Antennapedia to the basal transcriptional machinery. *Development* **135**, 1669-79.
Prochiantz, A. (2000). Messenger proteins: homeoproteins, TAT and others. *Curr Opin Cell Biol* **12**, 400-6.
Prochiantz, A. (2008). Protein and peptide transduction, twenty years later a happy birthday. *Adv Drug Deliv Rev* **60**, 448-51.
Prochiantz, A. and Joliot, A. (2003). Can transcription factors function as cell-cell signalling molecules? *Nat Rev Mol Cell Biol* **4**, 814-9.
Purcell, A. W., McCluskey, J. and Rossjohn, J. (2007). More than one reason to rethink the use of peptides in vaccine design. *Nat Rev Drug Discov* **6**, 404-14.
Qian, Y. Q., Billeter, M., Otting, G., Muller, M., Gehring, W. J. and Wuthrich, K. (1989). The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* **59**, 573-80.
Qian, Y. Q., Furukubo-Tokunaga, K., Resendez-Perez, D., Muller, M., Gehring, W. J. and Wuthrich, K. (1994a). Nuclear magnetic resonance solution structure of the fushi tarazu homeodomain from Drosophila and comparison with the Antennapedia homeodomain. *J Mol Biol* **238**, 333-45.
Qian, Y. Q., Resendez-Perez, D., Gehring, W. J. and Wuthrich, K. (1994b). The des(1-6)antennapedia homeodomain: comparison of the NMR solution structure and the DNA-binding affinity with the intact Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **91**, 4091-5.
Quiring, R., Walldorf, U., Kloter, U. and Gehring, W. J. (1994). Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans. *Science* **265**, 785-9.
Radtke, F., Wilson, A. and MacDonald, H. R. (2005). Notch signaling in hematopoiesis and lymphopoiesis: lessons from Drosophila. *Bioessays* **27**, 1117-28.
Rauer, B., Neumann, E., Widengren, J. and Rigler, R. (1996). Fluorescence correlation spectrometry of the interaction kinetics of tetramethylrhodamin alpha-bungarotoxin with Torpedo californica acetylcholine receptor. *Biophysical Chemistry* **58**, 3-12.
Rawat, V. P., Thoene, S., Naidu, V. M., Arseni, N., Heilmeier, B., Metzeler, K., Petropoulos, K., Deshpande, A., Quintanilla-Martinez, L., Bohlander, S. K. et al. (2008a). Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood* **111**, 309-19.

Rawat, V. P. S., Thoene, S., Naidu, V. M., Arseni, N., Heilmeier, B., Metzeler, K., Petropoulos, K., Deshpande, A., Quintanilla-Martinez, L., Bohlander, S. K. et al. (2008b). Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood* **111**, 309-319.

Reed, H. C., Hoare, T., Thomsen, S., Weaver, T. A., White, R. A., Akam, M. and Alonso, C. R. Alternative Splicing Modulates Ubx Protein Function in Drosophila melanogaster. *Genetics*.

Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M. and McGinnis, W. (1985). Homeo box genes of the Antennapedia and bithorax complexes of Drosophila. *Cell* **43**, 71-80.

Regulski, M., McGinnis, N., Chadwick, R. and McGinnis, W. (1987). Developmental and molecular analysis of Deformed; a homeotic gene controlling Drosophila head development. *EMBO J* **6**, 767-777.

Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. and Mann, R. S. (1997). Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell* **91**, 171-83.

Rigler, R. and Mets, U. (1992). Diffusion of single molecules through a Gaussian laser beam. *SPIE. Laser Spectroscopy of Biomolecules* **1921**, 239-248.

Rigler, R., Mets, U., Widengren, J. and Kask, P. (1993). Fluorescence Correlation Spectroscopy with High Count Rate and Low-Background - Analysis of Translational Diffusion. *European Biophysics Journal with Biophysics Letters* **22**, 169-175.

Rigler, R. and Widengren, J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy. *Bioscience* **3**, 180-183.

Rippe, K. (2000). Simultaneous binding of two DNA duplexes to the NtrC-enhancer complex studied by two-color fluorescence cross-correlation spectroscopy. *Biochemistry* **39**, 2131-9.

Rogers, B. T., Peterson, M. D. and Kaufman, T. C. (1997). Evolution of the insect body plan as revealed by the Sex combs reduced expression pattern. *Development* **124**, 149-57.

Rohs, R., West, S. M., Sosinsky, A., Liu, P., Mann, R. S. and Honig, B. (2009). The role of DNA shape in protein-DNA recognition. *Nature* **461**, 1248-53.

Rotem, S. and Mor, A. (2009). Antimicrobial peptide mimics for improved therapeutic properties. *Biochim Biophys Acta* **1788**, 1582-92.

Rousso, D. L., Gaber, Z. B., Wellik, D., Morrisey, E. E. and Novitch, B. G. (2008). Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. *Neuron* **59**, 226-40.

Rubin, G. M. and Lewis, E. B. (2000). A brief history of Drosophila's contributions to genome research. *Science* **287**, 2216-8.

Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. *Science* **218**, 348-53.

Rubinstein, M. and Niv, M. Y. (2009). Peptidic modulators of protein-protein interactions: progress and challenges in computational design. *Biopolymers* **91**, 505-13.

Rubinsztein, D. C. (2002). Lessons from animal models of Huntington's disease. *Trends Genet* **18**, 202-9.

Ryan, J. F., Mazza, M. E., Pang, K., Matus, D. Q., Baxevanis, A. D., Martindale, M. Q. and Finnerty, J. R. (2007). Pre-bilaterian origins of the Hox cluster and the

Hox code: evidence from the sea anemone, Nematostella vectensis. *PLoS One* **2**, e153.
Ryoo, H. D. and Mann, R. S. (1999). The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev* **13**, 1704-16.
Sanford, K. and Kumar, M. (2005). New proteins in a materials world. *Curr Opin Biotechnol* **16**, 416-21.
Saucedo, L. J. and Edgar, B. A. (2002). Why size matters: altering cell size. *Curr Opin Genet Dev* **12**, 565-71.
Schulz, B., Banuett, F., Dahl, M., Schlesinger, R., Schafer, W., Martin, T., Herskowitz, I. and Kahmann, R. (1990). The b alleles of U. maydis, whose combinations program pathogenic development, code for polypeptides containing a homeodomain-related motif. *Cell* **60**, 295-306.
Scott, M. P., Tamkun, J. W. and Hartzell, G. W., 3rd. (1989). The structure and function of the homeodomain. *Biochim Biophys Acta* **989**, 25-48.
Scott, M. P. and Weiner, A. J. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila. *Proc Natl Acad Sci U S A* **81**, 4115-9.
Seimiya, M., Ishiguro, H., Miura, K., Watanabe, Y. and Kurosawa, Y. (1994). Homeobox-containing genes in the most primitive metazoa, the sponges. *Eur J Biochem* **221**, 219-25.
Seimiya, M. and Kurosawa, Y. (1996). Kinetics of binding of Antp homeodomain to DNA analyzed by measurements of surface plasmon resonance. *FEBS Lett* **398**, 279-84.
Sevenich, F. W., Langowski, J., Weiss, V. and Rippe, K. (1998). DNA binding and oligomerization of NtrC studied by fluorescence anisotropy and fluorescence correlation spectroscopy. *Nucleic Acids Res* **26**, 1373-81.
Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E. and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* **22**, 1567-72.
Shen, W., Chrobak, D., Krishnan, K., Lawrence, H. J. and Largman, C. (2004a). HOXB6 protein is bound to CREB-binding protein and represses globin expression in a DNA binding-dependent, PBX interaction-independent process. *J Biol Chem* **279**, 39895-904.
Shen, W. F., Chrobak, D., Krishnan, K., Lawrence, H. J. and Largman, C. (2004b). HOXB6 protein is bound to CREB-binding protein and represses globin expression in a DNA binding-dependent, PBX interaction-independent process. *Journal of Biological Chemistry* **279**, 39895-39904.
Shen, W. F., Montgomery, J. C., Rozenfeld, S., Moskow, J. J., Lawrence, H. J., Buchberg, A. M. and Largman, C. (1997a). AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol Cell Biol* **17**, 6448-58.
Shen, W. F., Rozenfeld, S., Lawrence, H. J. and Largman, C. (1997b). The Abd-B-like Hox homeodomain proteins can be subdivided by the ability to form complexes with Pbx1a on a novel DNA target. *J Biol Chem* **272**, 8198-206.
Shepherd, J. C., McGinnis, W., Carrasco, A. E., De Robertis, E. M. and Gehring, W. J. (1984). Fly and frog homoeo domains show homologies with yeast mating type regulatory proteins. *Nature* **310**, 70-1.
Shera, E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A. and Soper, S. A. (1990). Detection of Single Fluorescent Molecules. *Chemical Physics Letters* **174**, 553-557.

Shi, X., Bai, S., Li, L. and Cao, X. (2001). Hoxa-9 represses transforming growth factor-beta-induced osteopontin gene transcription. *J Biol Chem* **276**, 850-5.
Shi, X., Yang, X., Chen, D., Chang, Z. and Cao, X. (1999). Smad1 interacts with homeobox DNA-binding proteins in bone morphogenetic protein signaling. *J Biol Chem* **274**, 13711-7.
Shimomura, O., Johnson, F. H. and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol* **59**, 223-39.
Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K. et al. (2000). Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. *Nat Genet* **25**, 302-5.
Shimura, H., Schlossmacher, M. G., Hattori, N., Frosch, M. P., Trockenbacher, A., Schneider, R., Mizuno, Y., Kosik, K. S. and Selkoe, D. J. (2001). Ubiquitination of a new form of alpha-synuclein by parkin from human brain: implications for Parkinson's disease. *Science* **293**, 263-9.
Shulman, J. M., Shulman, L. M., Weiner, W. J. and Feany, M. B. (2003). From fruit fly to bedside: translating lessons from Drosophila models of neurodegenerative disease. *Curr Opin Neurol* **16**, 443-9.
Shultz, J. W. and Regier, J. C. (2000). Phylogenetic analysis of arthropods using two nuclear protein-encoding genes supports a crustacean + hexapod clade. *Proc Biol Sci* **267**, 1011-9.
Siegel, P. M. and Massague, J. (2003). Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* **3**, 807-21.
Skouloudaki, K., Puetz, M., Simons, M., Courbard, J. R., Boehlke, C., Hartleben, B., Engel, C., Moeller, M. J., Englert, C., Bollig, F. et al. (2009). Scribble participates in Hippo signaling and is required for normal zebrafish pronephros development. *Proc Natl Acad Sci U S A* **106**, 8579-84.
Sommer, H., Beltran, J. P., Huijser, P., Pape, H., Lonnig, W. E., Saedler, H. and Schwarz-Sommer, Z. (1990). Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. *EMBO J* **9**, 605-13.
Sonnier, L., Le Pen, G., Hartmann, A., Bizot, J. C., Trovero, F., Krebs, M. O. and Prochiantz, A. (2007). Progressive loss of dopaminergic neurons in the ventral midbrain of adult mice heterozygote for Engrailed1. *J Neurosci* **27**, 1063-71.
Spradling, A. C. and Rubin, G. M. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. *Science* **218**, 341-7.
Stomel, J. M., Wilson, J. W., Leon, M. A., Stafford, P. and Chaput, J. C. (2009). A man-made ATP-binding protein evolved independent of nature causes abnormal growth in bacterial cells. *PLoS One* **4**, e7385.
Stroeher, V. L., Gaiser, J. C. and Garber, R. L. (1988). Alternative RNA splicing that is spatially regulated: generation of transcripts from the Antennapedia gene of Drosophila melanogaster with different protein-coding regions. *Mol Cell Biol* **8**, 4143-54.
Sulochana, K. N. and Ge, R. (2007). Developing antiangiogenic peptide drugs for angiogenesis-related diseases. *Curr Pharm Des* **13**, 2074-86.
Tapon, N., Harvey, K. F., Bell, D. W., Wahrer, D. C., Schiripo, T. A., Haber, D. A. and Hariharan, I. K. (2002). salvador Promotes both cell cycle exit and apoptosis in Drosophila and is mutated in human cancer cell lines. *Cell* **110**, 467-78.
Tolkunova, E. N., Fujioka, M., Kobayashi, M., Deka, D. and Jaynes, J. B. (1998). Two distinct types of repression domain in engrailed: one interacts with the groucho

corepressor and is preferentially active on integrated target genes. *Mol Cell Biol* **18**, 2804-14.
Tour, E., Hittinger, C. T. and McGinnis, W. (2005). Evolutionarily conserved domains required for activation and repression functions of the Drosophila Hox protein Ultrabithorax. *Development* **132**, 5271-81.
Van Auken, K., Weaver, D., Robertson, B., Sundaram, M., Saldi, T., Edgar, L., Elling, U., Lee, M., Boese, Q. and Wood, W. B. (2002). Roles of the Homothorax/Meis/Prep homolog UNC-62 and the Exd/Pbx homologs CEH-20 and CEH-40 in C. elegans embryogenesis. *Development* **129**, 5255-68.
Venken, K. J. and Bellen, H. J. (2007). Transgenesis upgrades for Drosophila melanogaster. *Development* **134**, 3571-84.
Vincent, J. P., Kassis, J. A. and O'Farrell, P. H. (1990). A synthetic homeodomain binding site acts as a cell type specific, promoter specific enhancer in Drosophila embryos. *EMBO J* **9**, 2573-8.
Vukojevic, V., Papadopoulos, D. K., Terenius, L., Gehring, W. J. and Rigler, R. (2010). Quantitative study of synthetic Hox transcription factor-DNA interactions in live cells. *Proc Natl Acad Sci U S A* **107**, 4093-8.
Vukojevic, V., Pramanik, A., Yakovleva, T., Rigler, R., Terenius, L. and Bakalkin, G. (2005). Study of molecular events in cells by fluorescence correlation spectroscopy. *Cell Mol Life Sci* **62**, 535-50.
Wagner-Bernholz, J. T., Wilson, C., Gibson, G., Schuh, R. and Gehring, W. J. (1991). Identification of target genes of the homeotic gene Antennapedia by enhancer detection. *Genes Dev* **5**, 2467-80.
Walsh, C. M. and Carroll, S. B. (2007). Collaboration between Smads and a Hox protein in target gene repression. *Development* **134**, 3585-92.
Wang, L. H., Chmelik, R. and Nirenberg, M. (2002). Sequence-specific DNA binding by the vnd/NK-2 homeodomain of Drosophila. *Proc Natl Acad Sci U S A* **99**, 12721-6.
Warrick, J. M., Chan, H. Y., Gray-Board, G. L., Chai, Y., Paulson, H. L. and Bonini, N. M. (1999). Suppression of polyglutamine-mediated neurodegeneration in Drosophila by the molecular chaperone HSP70. *Nat Genet* **23**, 425-8.
Wells, J. A. and McClendon, C. L. (2007). Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001-9.
Wiedenmann, J., Schenk, A., Rocker, C., Girod, A., Spindler, K. D. and Nienhaus, G. U. (2002). A far-red fluorescent protein with fast maturation and reduced oligomerization tendency from Entacmaea quadricolor (Anthozoa, Actinaria). *Proc Natl Acad Sci U S A* **99**, 11646-51.
Wieschaus, E. and Gehring, W. (1976). Clonal analysis of primordial disc cells in the early embryo of Drosophila melanogaster. *Dev Biol* **50**, 249-63.
Williams, T. M., Williams, M. E. and Innis, J. W. (2005). Range of HOX/TALE superclass associations and protein domain requirements for HOXA13:MEIS interaction. *Dev Biol* **277**, 457-71.
Wilson, C., Pearson, R. K., Bellen, H. J., O'Kane, C. J., Grossniklaus, U. and Gehring, W. J. (1989). P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. *Genes Dev* **3**, 1301-13.
Woerdeman, D. L., Veraverbeke, W. S., Parnas, R. S., Johnson, D., Delcour, J. A., Verpoest, I. and Plummer, C. J. (2004). Designing new materials from wheat protein. *Biomacromolecules* **5**, 1262-9.

Wu, J. and Cohen, S. M. (2000). Proximal distal axis formation in the Drosophila leg: distinct functions of teashirt and homothorax in the proximal leg. *Mech Dev* **94**, 47-56.
Xu, H., Frank, J., Trier, U., Hammer, S., Schroder, W., Behlke, J., Schafer-Korting, M., Holzwarth, J. F. and Saenger, W. (2001). Interaction of fluorescence labeled single-stranded DNA with hexameric DNA-helicase RepA: a photon and fluorescence correlation spectroscopy study. *Biochemistry* **40**, 7211-8.
Xu, T. and Rubin, G. M. (1993). Analysis of genetic mosaics in developing and adult Drosophila tissues. *Development* **117**, 1223-37.
Xu, T., Wang, W., Zhang, S., Stewart, R. A. and Yu, W. (1995). Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase. *Development* **121**, 1053-63.
Yakovleva, T., Pramanik, A., Kawasaki, T., Tan-No, K., Gileva, I., Lindegren, H., Langel, U., Ekstrom, T. J., Rigler, R., Terenius, L. et al. (2001). p53 Latency. C-terminal domain prevents binding of p53 core to target but not to nonspecific DNA sequences. *J Biol Chem* **276**, 15650-8.
Yallowitz, A. R., Gong, K. Q., Swinehart, I. T., Nelson, L. T. and Wellik, D. M. (2009). Non-homeodomain regions of Hox proteins mediate activation versus repression of Six2 via a single enhancer site in vivo. *Dev Biol* **335**, 156-65.
Yang, Y., Nishimura, I., Imai, Y., Takahashi, R. and Lu, B. (2003). Parkin suppresses dopaminergic neuron-selective neurotoxicity induced by Pael-R in Drosophila. *Neuron* **37**, 911-24.
Yao, J., Munson, K. M., Webb, W. W. and Lis, J. T. (2006). Dynamics of heat shock factor association with native gene loci in living cells. *Nature* **442**, 1050-3.
Yeh, B. J., Rutigliano, R. J., Deb, A., Bar-Sagi, D. and Lim, W. A. (2007). Rewiring cellular morphology pathways with synthetic guanine nucleotide exchange factors. *Nature* **447**, 596-600.
Yeh, H. C., Puleo, C. M., Lim, T. C., Ho, Y. P., Giza, P. E., Huang, R. C. and Wang, T. H. (2006). A microfluidic-FCS platform for investigation on the dissociation of Sp1-DNA complex by doxorubicin. *Nucleic Acids Res* **34**, e144.
Yusuf-Makagiansar, H., Anderson, M. E., Yakovleva, T. V., Murray, J. S. and Siahaan, T. J. (2002). Inhibition of LFA-1/ICAM-1 and VLA-4/VCAM-1 as a therapeutic approach to inflammation and autoimmune diseases. *Med Res Rev* **22**, 146-67.
Zeng, W., Andrew, D. J., Mathies, L. D., Horner, M. A. and Scott, M. P. (1993). Ectopic expression and function of the Antp and Scr homeotic genes: the N terminus of the homeodomain is critical to functional specificity. *Development* **118**, 339-52.
Zhang, J. and Nei, M. (1996). Evolution of Antennapedia-class homeobox genes. *Genetics* **142**, 295-303.
Zhang, Y. Q., Bailey, A. M., Matthies, H. J., Renden, R. B., Smith, M. A., Speese, S. D., Rubin, G. M. and Broadie, K. (2001). Drosophila fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* **107**, 591-603.
Zhao, J. J., Lazzarini, R. A. and Pick, L. (1996). Functional dissection of the mouse Hox-a5 gene. *EMBO J* **15**, 1313-22.
Zorrilla, S., Lillo, M. P., Chaix, D., Margeat, E., Royer, C. A. and Declerck, N. (2008a). Investigating transcriptional regulation by fluorescence spectroscopy, from traditional methods to state-of-the-art single-molecule approaches. *Ann N Y Acad Sci* **1130**, 44-51.

Zorrilla, S., Ortega, A., Chaix, D., Alfonso, C., Rivas, G., Aymerich, S., Lillo, M. P., Declerck, N. and Royer, C. A. (2008b). Characterization of the control catabolite protein of gluconeogenic genes repressor by fluorescence cross-correlation spectroscopy and other biophysical approaches. *Biophys J* **95**, 4403-15.

CURRICULUM VITAE

Name: Dimitrios Papadopoulos

Date of birth: 12.07.1983

Address: Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056

E-mail: d.papadopoulos@unibas.ch

Telephone: +41612671616

Current position: Post-doctoral scientist in the laboratory of Prof. Walter J. Gehring in the Department of Cell Biology of the University of Basel, Switzerland

Education

15.10.2006-14.06.2010 Ph.D. Genetics, University of Basel, Switzerland

Supervisor: Prof. Walter J. Gehring. Doctorate thesis: “Function of synthetic Hox transcription factors in vivo and the quantitative study of their molecular interactions with DNA at the single-molecule level in live cells”.

Evaluation: *Summa Cum Laude*, grade 6 (max. 6)

2001-2006 Diploma in Biology, University of Crete, Greece

Supervisor Prof. Kyriakos Kotzabasis. Diploma thesis: “The effect of Ca^{2+} and polyamines on the molecular structure and function of the photosynthetic apparatus”. Evaluation: 10 (max. 10).

2004-2005 Socrates-Erasmus student exchange program (one semester), University of Valencia, Spain

Languages

Greek (mother tongue), English (fluent level), German (fluent level), Spanish (fluent level)

Publications in Peer Reviewed International Journals

1) **Papadopoulos, D.K.**, Vukojevic, V., Adachi, Y., Terenius, L., Rigler, R., and Gehring, W.J. 2010. Function and specificity of synthetic Hox transcription factors in vivo. *Proc Natl Acad Sci U S A* **107**(9): 4087-4092.

2) Vukojevic, V., **Papadopoulos, D.K.**, Terenius, L., Gehring, W.J., and Rigler, R. 2010. Quantitative study of synthetic Hox transcription factor-DNA interactions in live cells. *Proc Natl Acad Sci U S A* **107**(9): 4093-4098.

Manuscripts

1) **Papadopoulos, D.K.**, Reséndez-Pérez, D., Cárdenas-Chávez, D.L., Villanueva-Segura, K., Canales del Castillo, R., Felix, D.A., Fünfschilling, R., and Gehring, W.J. 2010. Functional synthetic Antennapedia genes: dual role of the YPWM motif in transcriptional activation and repression.

2) Vukojević, V., **Papadopoulos, D.K.**, Terenius, L., Gehring, W.J., Rigler, R. 2010. Quantitative study of synthetic Hox transcription factor quest for specific targets in live cells.

Additional research training

01.12.2003-31.12.2003

Summer school at the National Agriculture Research Foundation (N.Ag.Re.F.) in Heraklion, Crete. Supervisor Dr. Andreas Doulis. Project: “Identification of pathogenic strains of the fungus *Fusarium oxysporum*, using the R.A.P.D. (Random Amplified polymorphic DNA) technique”.

01.07.2004 – 15.08.2004

Summer school at the laboratory of Plant Biochemistry and Physiology at the Department of Biology of the University of Crete. Supervisor Prof. Kalliopi Roubelakis-Angellakis. Project: “Quantitative analysis of the induction of the alternative oxidase gene (AOX) using the Western plot technique on single and double antisense strains of *Nicotiana tabacum*, after treatment with different concentrations of NaCl”.

01.05.2006 – 31.07.2006

Practical course at the laboratory of Molecular Biology and Biotechnology of plants of the University of Freiburg. Supervisor Dr. Minako Ueda. Project: "*Cis*-element analysis of the WOX2-promoter in Arabidopsis thaliana".

Attended conferences

1) **1st International Congress of Biotechnology** Athens, Greece (10-12.07.2004).
2) **14th F.E.S.P.B.** (Federation of European Societies of Plant Biologists) **Congress** Krakow, Poland (23-27.08.2004).
3) **2nd International Congress of Biotechnology** Athens, Greece (01-03.07.2005).
4) **Spetses International Summer School 2007: Molecular mechanisms of regeneration.** From stem cells to organ development, function and regeneration (31.08-09.09.2007).
5) **3rd Advanced Live Cell Microscopy Workshop** Madrid, Spain (04-06.10.2007).
6) **Biozentrum Symposium** Basel, Switzerland (25.01.2008). Poster presentation *In vivo* DNA-transcription factor interactions visualized by FCS".
7) **Nobel Symposia "Single Molecule Spectroscopy in Chemistry, Physics and Biology"** Sånga-Säby, Sweden (01-06.06.2008). Poster presentation: "Quantitative Confocal Laser Scanning Microscopy imaging of transcriptional activity in live cells".
8) **Frontiers in the New Biology: WCN Symposium on Functional Genomics 2008** Uppsala, Sweden (29.09.2009). Poster presentation: "Quantitative Confocal Laser Scanning Microscopy imaging of transcriptional activity in live cells".
9) **Biozentrum Symposium** Basel, Switzerland (18.12.2008). Poster presentation: "Live cell transcriptional dynamics of a functional synthetic Hox gene".
10) **Cells into Organs Symposium: Tissue Specification and Organogenesis** Lisbon, Portugal (04-06.02.2009). Poster presentation: "Live cell transcriptional dynamics of a functional synthetic Hox gene".
11) **12th Carl Zeiss sponsored workshop on Fluorescence Correlation Spectroscopy and related methods** Cargese, France (12-16.10.2009). Poster presentation: "Live cell transcriptional dynamics of a functional synthetic Hox gene" and talk: "Functional synthetic Hox genes: imaging the activity and quantifying transcription factor-DNA interactions in live cells" (Vukojević V, Papadopoulos DK, Terenius L, Gehring WJ and Rigler R).