DECODING THE LIM DEVELOPMENT CODE

GORDON N. GILL, M.D.

LA JOLLA, CALIFORNIA

ABSTRACT

During development a vast number of distinct cell types arise from dividing progenitor cells. Concentration gradients of ligands that act via cell surface receptors signal transcriptional regulators that repress and activate particular genes. LIM homeodomain proteins are an important class of transcriptional regulators that direct cell fate. Although in *C. elegans* only a single LIM homeodomain protein is expressed in a particular cell type, in vertebrates combinations of LIM homeodomain proteins are expressed in cells that determine cell fates. We have investigated the molecular basis of the LIM domain "code" that determines cell fates such as wing formation in *Drosophilia* and motor neuron formation in chicks. The basic code is a homotetramer of 2 LIM homeodomain proteins bridged by the adaptor protein, nuclear LIM interactor (NLI). A more complex molecular language consisting of a hexamer complex involving NLI and 2 LIM homeodomain proteins, Lhx3 and Isl1 determines ventral motor neuron formation. The same molecular "words" adopt different meanings depending on the context of expression of other molecular "words."

INTRODUCTION

Different cell types arise from dividing progenitor cells in response to concentration gradients of ligands that act via cell surface receptors and to a combination of transcriptional regulators that repress or activate particular genes. In vertebrates 12 LIM homeodomain genes and 4 nuclear LIM only proteins have been identified. Deletion or mutation of individual LIM homeodomain genes have profound effects illustrating the essential role these proteins play in developmental processes. Tanabe and Jessell (1) first described the regional pattern of expression of several LIM homeodomain genes in the developing spinal cord and proposed that distinct motor neurons develop in response to the repertoire of LIM homeodomain proteins expressed. These observations were in basis of the LIM "code" hypothesis, which states that the particular combination of LIM genes expressed specifies distinct cell fates. Because LIM homeodomain genes are expressed in different

combinations, this raises the general question of how specificity is achieved using a limited number of factors. A related question is how does the information provided by one molecule result in different outcomes when expressed in the context of other molecules.

LIM domains are 50 amino acid 2 Zn^{2+} finger protein modules that are found in a variety of cell proteins. The word LIM drives from the first 3 proteins noted to have LIM domains, Lin 11, Isl1, and Mec 3 (2). NMR structural analysis indicates LIM domains consist of a series of anti-parallel β sheets, a hydrophobic core and 2 tetrahedrally coordinated Zn^{2+} molecules (3). There are two major groups of LIM proteins: nuclear and cytoplasmic. Nuclear LIM homeodomain proteins contain a pair of closely spaced LIM domains located NH_2—terminal to a homeodomain. Nuclear LIM only proteins (LMO), that were initially discovered as oncogenes in human T cell leukemia (4), consist of a pair of LIM domains without other sequence information. Cytoplasmic LIM domain proteins contain variable numbers of LIM domains without or with other functional domains. Most if not all cytoplasmic LIM proteins function in cytoskeleton regulation (5).

The key to deciphering the molecular basis of the LIM code was the discovery of the Nuclear Lim Interactor, NLI, also called LIM domain binder, Ldb, cofactor for LIM domains, CLIM, or the Drosophilia ortholog, Chip (6–9). NLI binds the pair of LIM domains of all LIM homeodomain and nuclear LMO proteins with high affinity but does not recognize cytoplasmic LIM domains. Functional analysis of NLI revealed that it consists of an NH_2 terminal dimerization domain (DD), a nuclear localization sequence and a COOH-terminal LIM interaction domain (LID) (10). NLI is thus an adapter for nuclear LIM proteins mediating formation of complexes containing various LIM homeodomain and LMO proteins.

NLI MEDIATES INTERACTIONS AMONG NUCLEAR LIM PROTEINS

NLI was identified in an expression overlay screen using an LMO2 probe (6). NLI recognizes the paired LIM domains with much higher affinity than either LIM domain alone. The 38 amino acid LID of NLI consists of NH_2-terminal hydrophobic and COOH-terminal helical regions both of which are required for the high affinity interaction with the paired LIM domains (Figure 1). The NH_2 terminal 200 amino acids of NLI dimerize mediating formation of strong homodimers (10). NLI thus forms tetrameric complexes consisting of two NLI molecules complexed with two LIM homeodomain proteins (Figure 2). With the

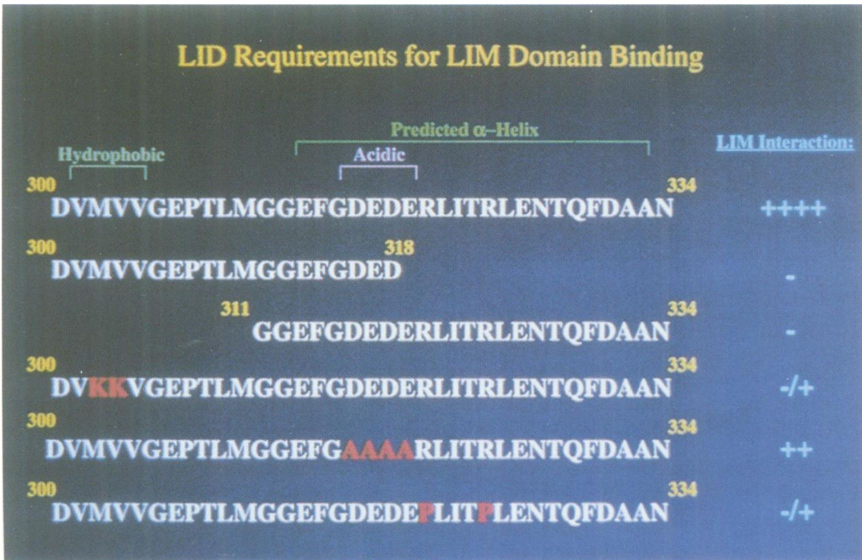

FIG. 1. Amino acid sequence and functional analysis of the LIM interaction domain of NLI. The interaction of wild type and mutant LIDs with paired nuclear LIM domains was quantitated in coimmunoprecipitation experiments and is scored on a scale of 0 to 4+. NH_2- or COOH-terminal deletions disrupt binding. Mutation of the NH_2-terminal hydrophobic region or disruption of the COOH-terminal helical domain with proline substitutions also abrogates binding.

exception of the direct interaction of Lhx3 with Isl, LIM homeodomain proteins do not directly interact but require NLI as a bridge. NLI mediates formation of both homodimeric and heterodimeric LIM homeodomain complexes (11). Although LMO proteins may form complexes with basic helix-loop-helix proteins (12), LMOs most likely bind one LID to disrupt a binary LIM homeodomain complex thus stalling development.

A HOMOTETRAMERIC LIM CODE DIRECTS DEVELOPMENT OF DROSOPHILIA WINGS AND INTERNEURONS

The LIM homeodomain protein apterous (*ap*) is required for dorsal wing and for interneuron development (13,14). In the absence of *ap* wings do not form, nor do they form when the LIM domains of *ap* are deleted. (Figure 3A, upper 2 panels). When mutant Chip/NLI that lack either the dimerization or LID domains and thus interfere with the bridging function of the endogenous protein are expressed, wings also fail to form (Figure 3A, lower 2 panels) (15).

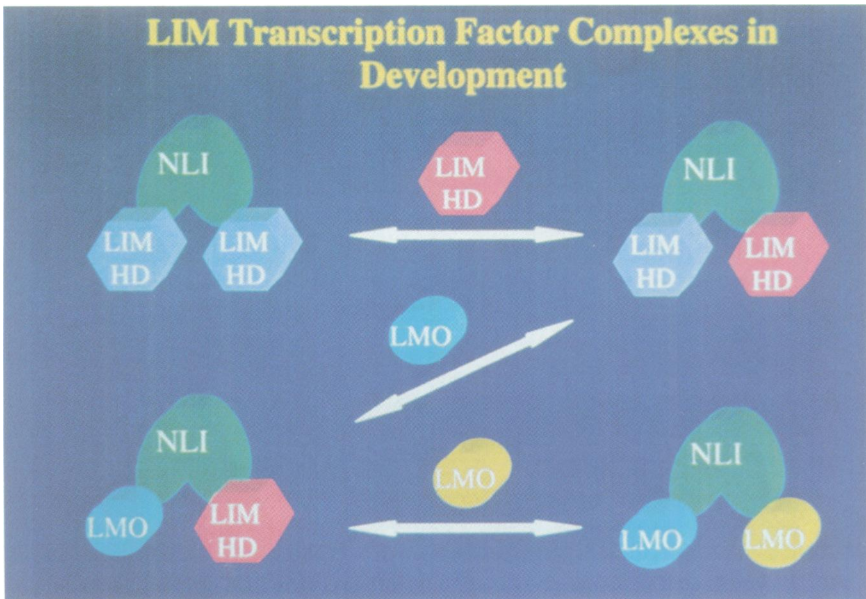

FIG. 2. Some LIM transcription factor complexes in development. Dimeric NLI molecules bind 2 LIM homeodomain proteins to create 2NLI · 2LIM homeodomain protein tetramers. These can consist of similar or different LIM homeodomain proteins. LMO proteins displace one or both LIM homeodomains disrupting the tetrameric code.

A prediction of a homotetramer code is that proper complex formation depends on appropriate stoichiometry of NLI and its LIM homeodomain partner. Biochemically excess NLI distributes the LIM homeodomain proteins into trimers as well as tetramers. Proper stoichiometry can be reestablished by lowering NLI or increasing LIM homeodomain concentrations. As shown in Figure 4, middle panel, overexpression of NLI/Chip severely disturbs wing formation; increased expression of *ap*, reestablishes stoichiometry and rescues this defect (Figure 4, right panel). A critical test of the tetramer model was to create a chimeric molecule in which the dimer domain of NLI/Chip was fused to the homeodomain of *ap*, bypassing the LID/LIM interactions. This chimeric molecule rescued wing formation in flies lacking both NLI/Chip and *ap* (Figure 3B). The programs of wing and *ap* interneuron formation are thus directed by homotetrameric complexes composed of 2NLI · 2LIM homeodomain proteins (15,16). Development of *ap* interneurons is also directed by this homotetrameric code (17).

Expression of LM02 in transgenic mice reproduces the human T cell leukemia associated with translocation and aberrant expression of LM02 t(11; 14)(p15; q11) (18,19). Misexpression of LMO1 or LMO2 in

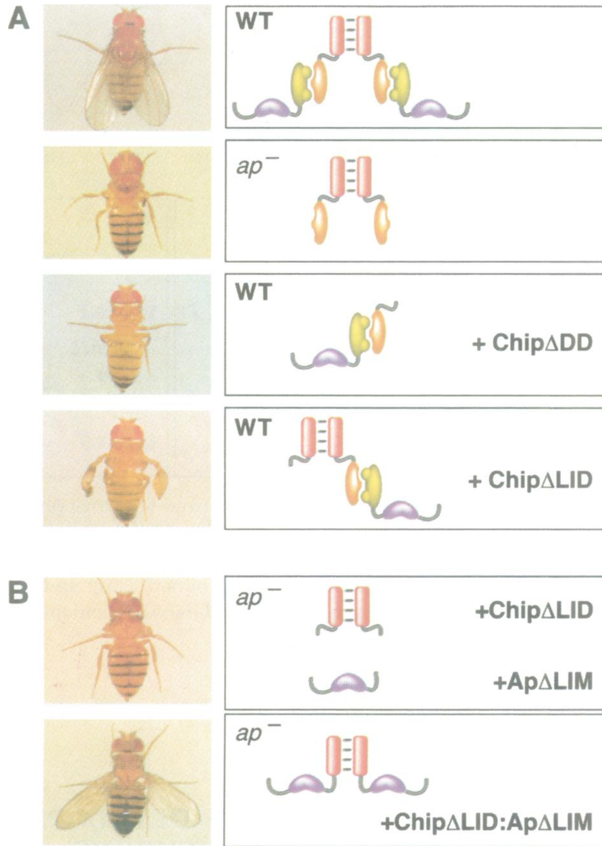

FIG. 3. A homotetrameric 2NLI/Chip · *ap* complex directs Drosophilia wing development. A. Upper panel: single copy of wild-type *ap* with normal wing formation. Second panel: absence of wing formation in *ap* −/−. Third panel: absence of wings when Chip lacking the dimer domain is expressed resulting in the LID sequestering *ap* from endogenous Chip. Fourth panel: aberrant wing formation when Chip lacking the LID is expressed resulting in aberrant Chip dimer formation. B. Upper panel: In the absence of *ap*, expression of Chip without the LID plus *ap* without its LIM domains fail to restore wing formation. Lower panel: Rescue of wing formation by a chimeric molecule consisting of the dimer domain of Chip fused directly to the homeodomain of *ap*. Modified from data of Van Meyel, et al. (15).

mouse thymus results in disease in ~9 months with an immature phenotype ($CD4^-$ $CD8^-$ or $CD4^+$ $CD8^+$). This pattern of disease is compatible with aberrant development of T cell precursors. In *Drosophilia*, wing development depends on expression of the LIM homeodomain protein *ap*; the developmental program is terminated by induc-

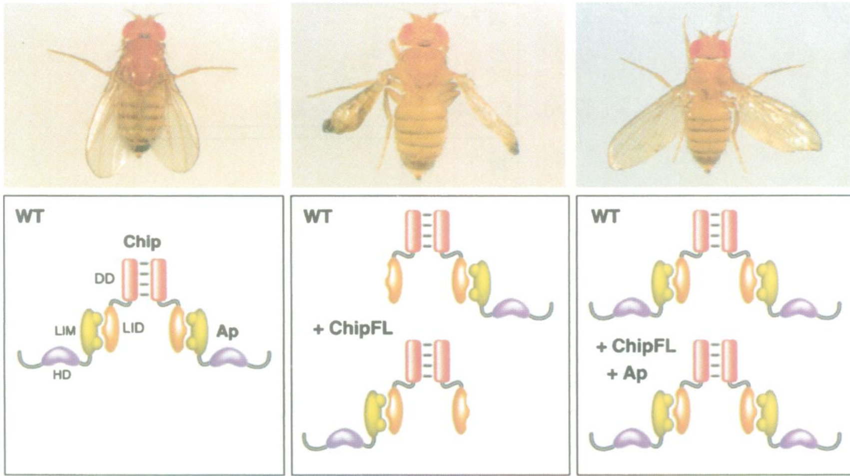

FIG. 4. Proper stoichiometry of Chip and *ap* is required for Drosophilia wing formation. Left panel: Normal wing formation in *ap* +/− flies. Middle panel: Defective wing formation with over expression of Chip. Excessive chip partitions *ap* into trimers. Right panel: Rescue of wing development by over expression of both Chip and *ap*. Overexpression of *ap* in a Chip overexpression background re-establishes stoichiometry appropriate for formation of tetrameric complexes.

tion of dLMO that presumably disrupts tetrameric 2 Chip · 2 *ap* complexes (15,16).

A HEXAMERIC LIM CODE DIRECTS DEVELOPMENT OF VERTEBRATE MOTOR NEURONS

V_2 interneurons (IN) and motor neurons (MN) in the ventral spinal cord of vertebrates participate in neuronal circuits for the coordination of locomotor activity (20). These two classes of cells develop adjacent to one another in response to the graded inductive activity of the ligand sonic hedgehog and a program of specific transcription factors (21). Both V_2 IN and ventral MN express the LIM homeodomain protein Lhx3 and the closely related Lhx4 (22) but only MN express Isl1. Thus the LIM code for V2 IN is Lhx3 and for ventral motor neurons is Lhx3 + Isl1 (Figure 5). Deletion of Lhx3/Lhx4 abolishes V_2 IN formation and MN fail to properly develop (23). Deletion of Isl1 disrupts MN formation altogether (24).

V_2 IN form in response to a tetrameric 2NLI · 2Lhx3 LIM code (25). These same complexes would be expected to form in MN, generating neurons with hybrid properties. Such does not occur nor do dorsally projecting motor neurons that express Isl1 only form in the ventral

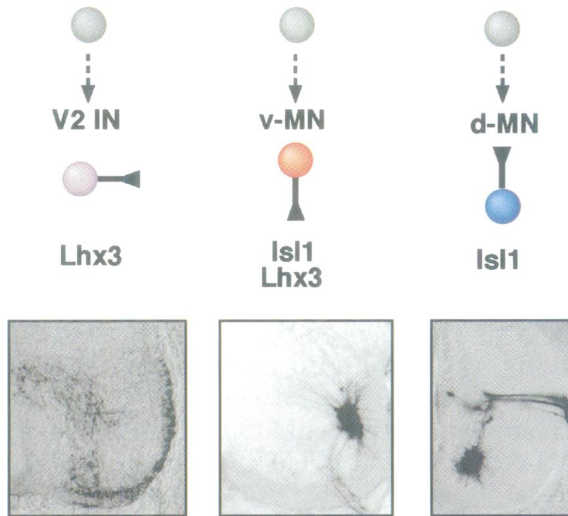

FIG. 5. The LIM expression code for spinal neurons. V2 interneurons express the LIM homeodomain protein Lhx3; ventrally projecting motor neurons from the lateral column express Lhx3 and Isl1; dorsally projecting motor neurons from the medial column express Isl1. Courtesy of S. L. Pfaff.

area. Strict segregation of cell fate is achieved by two distinct types of protein interactions involving the LIM domains of Lhx3 that convert it from a V_2 IN promoting factor to a MN promoting factor.

In ovo electroporation was used to introduce expressor plasmids into one side of the developing chick neural tube and development of specific cell types were scored by immunostaining for cell type specific markers: Chx 10, a paired-like homeodomain protein, for V_2 IN and HB9, a divergent homeodomain protein for MN. Ectopic expression of Lhx3 led to expanded development of V_2 IN while expression of Lhx3 and Isl1 led to expanded development of ventral MN (25). Formation of V_2 IN required both the LIM domains and the homeodomain of Lhx3. V_2 IN development was blocked by expression of mutant NLIs lacking either the dimer or LID domains or expression of the competitor LMO4 but could be reproduced by expression of a chimeric NLI-homeodomain fusion protein that contained the dimer domain of NLI attached directly to the homeodomain of Lhx3. V_2 IN thus develops from the use of a homotetrameric 2NLI · 2Lhx3 LIM code analogous to the Chip · *ap* code in Drosophilia.

In MN possible complexes include homotetramers of NLI with Lhx3 or with Isl1 or heterotetramers between an NLI dimer and both Lhx3 and Isl1. However, tetrameric complexes did not trigger MN develop-

ment. This was shown most clearly by coexpressing 2 chimeric molecules: one consisting of the dimer domain of NLI fused to the homeodomain of Isl1 and one consisting of the dimer domain of NLI fused to the homeodomain of Lhx3. These analogs which biochemically formed complexes in vitro failed to generate MN in vivo.

Previous studies had shown that the LIM domains of Lhx3 but not those of other LIM homeodomain or LMO proteins recognized not only the LID of NLI but also the homeodomain/Isl-specific domain of Isl1 and Isl2 (11). This unique property of the LIM domains of Lhx3 leads to the formation of hexameric complexes consisting of 2NLI · 2Isl1 · 2Lhx3 molecules that specify ventral MN development (Figure 6).

Experimentally both the LIM domains and homeodomains of Isl1 and of Lhx3 are required for MN development. The LIM domains of another LIM homeodomain protein Lhx1 could not substitute for those of Lhx3 in MN development although these could substitute in Lhx3 directed V_2 IN formation, since both LIM domain pairs recognize NLI. Evidence for the hexamer model came from expressing a chimeric NLI dimer domain fused to the homeodomain of Isl1 as an analog of the 2NLI · 2Isl1 complex plus co-expressing Lhx3 and observing MN formation. Because the hexamer complex involves the LIM domains of Lhx3 binding to the Isl-specific domain, coexpression of LMO4 or the LID of NLI do not disrupt MN formation. A single triple fusion consisting of the dimer domain of NLI fused to the homeodomains of both Isl1 and Lhx3 initiated MN programs.

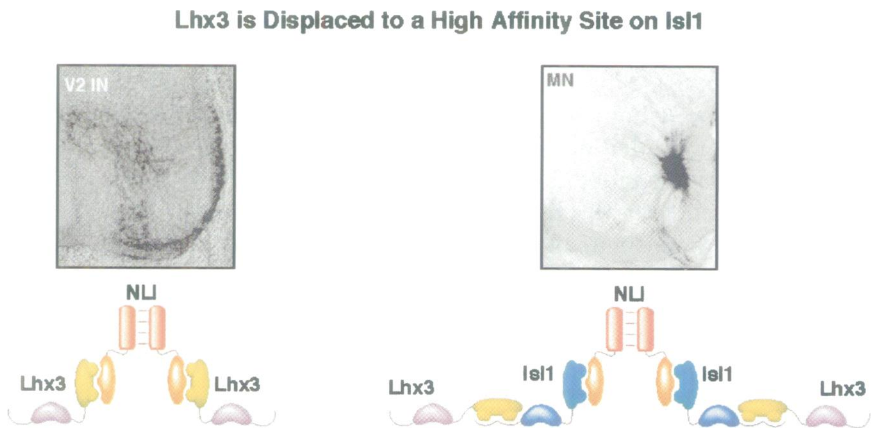

FIG. 6. Biochemical mechanism through which Lhx 3 directs V2 IN and MN generation. V2 IN are determined by a tetrameric 2NLI · 2Lhx3 LIM code. In MN Isl1 displaces Lhx3 from NLI to the Isl specific domain so that hexameric 2NLI · 2Isl 1 · 2Lhx 3 complexes form. Courtesy of S. L. Pfaff.

Isl1 thus converts Lhx3 from a V_2 IN factor to a MN factor by displacing it from NLI to interact directly with Isl1. When Lhx3 and Isl1 are coexpressed, hexamers form at the expense of tetramers resulting in a specific cell fate. The circuitry yielding highly specific cell fates is reinforced. HB9, whose transcription is enhanced by the hexameric complex, functions as a repressor of V_2 IN formation and as an enhancer of MN formation (26,27).

There are undoubtedly additional levels of complexity including interactions of LIM homeodomain proteins with other transcription factors and interaction with RLIM, a ubiquitin ligase that may facilitate turnover of LIM homeodomain proteins (28). Not yet addressed are questions of how assembly of tetrameric or hexameric LIM homeodomain complexes sited on promoters affect the core transcriptional machinery. What is clear is that LIM molecular "codes" initiate specific gene transcription programs that direct late cell fate decisions. Detailed knowledge of developmental programs is crucial to development of therapies based on stem cell differentiation (29).

ACKNOWLEDGMENTS

I am grateful to my colleagues whose work is described, especially to Linda Jurata who discovered NLI and its biochemical properties, to Don Van Meyel and John Thomas who translated biochemistry into genetic proof in *Drosophilia*, and to Josh Thaler, Soo-Kyung Lee and Sam Pfaff whose deep insight into neurobiology exposed the molecular underpinnings of the LIM code for motor neuron development. This work was supported by NIH grant DK13149.

REFERENCES

1. Tanabe Y, Jessell TM. Diversity and pattern in the developing spinal cord. *Science* 1996;274:1115–1123.
2. Freyd G, Kim SK, Horvitz HR. Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene lin-11. *Nature* 1990;344: 876–879.
3. Perez-Alvarado GC, Miles C, Michelsen JW, Louis HA, Winge DR, Beckerle MC, Summers MF. Structure of the carboxy-terminal LIM domain from the cysteine rich protein CRP. *Nat. Struct. Biol.* 1994;1:388–398.
4. Boehm T, Foroni L, Kaneko Y, Perutz MF, Rabbitts TH. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11 p. 15 and 11 p. 13. *Proc. Natl. Acad. Sci. USA* 1991;88:4367–4371.
5. Edwards DC, Sanders LC, Bokoch GM, Gill GN. Coupling of Rac/Cdc42 GTPase signaling to actin cytoskeletal dynamics by Pak activation of LIM kinase. *Nature Cell Biology* 1999;1:253–259.
6. Jurata LW, Kenny DA, Gill GN. Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development. *Proc. Natl. Acad. Sci. USA* 1996;93:11693–11698.

7. Agulnick AD, Taira M, Breen JJ, Tanaka T, Dawid IB, Westphal H. Interactions of the LIM domain binding factor Ldb1 with LIM homeodomain proteins. *Nature* 1996;384:270–272.
8. Bach I, Carriere C, Ostendorff HP, Andersen B, Rosenfeld MG. A family of LIM domain-associated cofactors confer transcriptional synergism between LIM and Otx homeodomain proteins. *Genes Dev.* 1997;11:1370–1380.
9. Morcello P, Rosen C, Baylies MK, Dorsett D. Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in *Drosophilia. Genes Dev.* 1997;11:2729–2740.
10. Jurata LW, Gill GN. Functional analysis of the nuclear LIM domain interactor, NLI. *Mol. Cell Biol.* 1997;17:5688–5698.
11. Jurata LW, Pfaff SL, Gill GN. The nuclear LIM interactor NLI mediates homo- and hetero dimerization of LIM domain transcription factors. *J. Biol. Chem.* 1998:273;3152–3157.
12. Wadman I, Li J, Bash RO, Forster A, Osada H, Rabbits TH, Baer R. Specific in vivo association between bHLH and LIM proteins implicated in human T cell leukemia. *Embo J* 1994;13:4831–4839.
13. Cohen B, Mcguffin ME, Pfeifle C, Segal D, Cohen SM. Apterous, a gene required for imaginal disk development in *Drosophilia* encodes a member of the LIM family of developmental regulatory proteins. *Genes Dev.* 1992;6:715–729.
14. Lundgren SE, Callahan CA, Thor S, Thomas JB. Control of neuronal pathway selection by the *Drosophilia* LIM homeodomain gene *apterous. Develop.* 1995;121:1769–1773.
15. Van Meyel DJ, O'Keffe DD, Jurata LW, Thor S, Gill GN, Thomas JB. Chip and apterous physically interact to form a functional complex during *Drosophilia* development. *Mol. Cell* 1999;4:259–265.
16. Milan M, Cohen SM. Regulation of LIM homeodomain activity in vivo: a tetramer of dLDB and apterous confers activity and capacity for regulation by dLMO. *Mol. Cell* 1999;4:267–273.
17. Van Meyel DJ, O'Keffe DD, Thor S, Jurata LW, Gill GN, Thomas JB. Chip is an essential cofactor for apterous in the regulation of axon guidance in *Drosophilia. Development* 2000;127:1823–1831.
18. McGuire EA, Rintoul CE, Solar GM, Korsmeyer SJ. Thymic over expression of Ttg-1 in transgenic mice results in T-cell acute lymphoblastic leukemia/lymphoma. *Mol. Cell Biol.* 1992;12:4186–4196.
19. Fisch P, Boem T, Lavenir I, Larson T, Arno J, Forster A, Rabbitts TH. T-cell acute lymphoblastic lymphoma induced in transgenic mice by RBTN1 and RBTN2 LIM-domain genes. *Oncogene* 1992;7:2389–2397.
20. Lee SK, Pfaff SL. Transcriptional networks regulating neuronal identity in the developing spinal cord. *Nat. Neurosci.* 2001;4 Suppl. 1:1183–1191.
21. Jessell TM. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat. Rev. Genet.* 2000;1:20–29.
22. Marquardt T, Pfaff SL. Cracking the transcriptional code for cell specification in the neural tube. *Cell* 2001;106:651–654.
23. Sharma K, Sheng HZ, Lettieri K, Li H, Karavanov A, Potter A, Westphal H, Pfaff SL. LIM homeodomain factors Lhx3 and Lhx4 assign subtype identities for motor neurons. *Cell* 1998;95:817–828.
24. Pfaff SL, Mendelsohn M, Steward CL, Edlund T, Jessell TM. Requirement for LIM homeobox gene Isl1 in motor neuron generation reveals a motor neuron-dependent step in interneuron development. *Cell* 1996;84:309–320.
25. Thaler JP, Lee S-K, Jurata LW, Gill GN, Pfaff SL. LIM factor Lhx3 contributes to the

specification of motor neuron and interneuron identity through cell type-specific protein-protein interactions. *Cell* 2002;110:1–20.
26. Arber S, Han B, Mendelsohn M, Smith M, Jessell TM, Sockanathan S. Requirement for the homeobox gene HB9 in the consolidation of motor neuron identity. *Neuron* 1999;23:675–687.
27. Thaler J, Harrison K, Sharma K, Lettieri K, Kehri J, Pfaff SL. Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. *Neuron* 1999;23:675–687.
28. Ostendorff HP, Peirano RI, Peters MA, Schluter A, Bossenz M, Scheffner M, Back I. Ubiquitin-dependent cofactor exchange on LIM homeodomain transcription factors. *Nature* 2002;416:99–103.
29. Wichterie H, Lieberam I, Porter JA, Jessell TM. Directed differentiation of embryonic stem cells into motor neurons. *Cell* 2002;110:385–397.

DISCUSSION

Wasserman, La Jolla: Gordon, could the ability to introduce these proteins in later life, prove to be a potential target for the correction of loss of motor neurons due to trauma let's say?

Gill, La Jolla: One hesitates to be an optimist, especially in public, but in fact, by understanding the kind of molecular codes that cause final cell differentiation, one hopes that that's exactly what one can do. There's a recent paper in Cell from Thomas Jessel's laboratory in which they took ES cells, programmed them in vitro and then put them into the chick spinal chord where they developed into motor neurons that made acetylcholine synapses with muscle. Now that's a bit of an artificial system but it's a real, hopeful kind of experiment for motor neuron generation. A simpler experiment that is being done in the Pfaff lab, is to introduce the set of molecules that I've talked to you about directly into mouse stem cells. You can essentially convert 100% of the stem cells into motor neurons as assessed by differentiation markers. So I think that by understanding the kind of molecules that drive differentiation, you do have the hope that you could engineer precursor cells into the correct final pathway.