A Structural Basis for the Regulation of the LIM-Homeodomain Protein Islet 1 (Isl1) by Intra- and Intermolecular Interactions*

Morgan S. Gadd 1, David A. Jacques 2, Ivan Nisevic 1, Vanessa J. Craig 1, Ann H. Kwan, J. Mitchell Guss, and Jacqueline M. Matthews 1

From the School of Molecular Bioscience, Building G08, University of Sydney, New South Wales 2006, Australia

Islet 1 (Isl1) is a transcription factor of the LIM-homeodomain (LIM-HD) protein family and is essential for many developmental processes. LIM-HD proteins all contain two protein-interacting LIM domains, a DNA-binding homeodomain (HD), and a C-terminal region. In Isl1, the C-terminal region also contains the LIM homeobox 3 (Lhx3)-binding domain (LBD), which interacts with the LIM domains of Lhx3. The LIM domains of Isl1 have been implicated in inhibition of DNA binding potentially through an intramolecular interaction with or close to the HD. Here we investigate the LBD as a candidate intramolecular interaction domain. Competitive yeast-two hybrid experiments indicate that the LIM domains and LBD from Isl1 can interact with apparently low affinity, consistent with no detection of an intermolecular interaction in the same system. Nuclear magnetic resonance studies show that the interaction is specific, whereas substitution of the LBD with peptides of the same amino acid composition but different sequence is not specific. We solved the crystal structure of a similar but higher affinity interaction domain from the LIM-HD cofactor protein Isl1-binding domain protein 1 (Ldb1) and used these coordinates to generate a homology model of the intramolecular interaction that indicates poorer complementarity for the weak intramolecular interaction. The intramolecular interaction in Isl1 may provide protection against aggregation, minimize unproductive DNA binding, and facilitate cofactor exchange within the cell.

LIM-homeodomain (LIM-HD)5 proteins are transcription factors that are critical in the development of many cell types and tissues. Mammals carry 12 LIM-HD genes, which exist as six pairs of closely related paralogues in each species: Isl1/Isl2, Lhx1/Lhx5, Lhx2/Lhx9, Lhx3/Lhx4, Lhx6/Lhx8, and Lmx1a/Lmx1b. LIM-HD proteins and their binding partners are expressed in overlapping but distinct patterns to specify cell types or tissues at various locations and stages of development (e.g. Refs. 1 and 2). These properties have led to the concept of the “LIM code” in which LIM-HD proteins act in a combinatorial manner to regulate tissue patterning at the level of transcription (3, 4).

All LIM-HD proteins contain two tandemly arrayed LIM domains at or near the N terminus that mediate protein-protein interactions followed by a homeodomain that binds DNA and a C terminus that is predicted to be unstructured (e.g. see Fig. 1). The activity of LIM-HD proteins is dependent on the essential cofactor LIM domain-binding protein 1 (Ldb1; for a review, see Ref. 5). Ldb1 contains an N-terminal self-association domain (6) that forms trimers in vitro (7) and mediates long range chromatin interactions (8–10). It also contains a ~30-residue LIM interaction domain (LID) that binds to the LIM domains of all LIM-HD and closely related LIM-only (LMO) proteins (6, 11–14). Ldb1LID binds as an extended peptide across both LIM domains of its binding partners (13, 15, 16). Isl1 and Isl2 each contain a LID-like sequence in their C-terminal regions, the Lhx3-binding domain (LBD), which binds the LIM domains Lhx3 and Lhx4 (15, 17). These four proteins are co-expressed in and are necessary for the formation of motor neurons in the ventral neural cord (2, 18–20). By binding the LIM domains of Lhx3 and Lhx4 (Lhx3/4Lint1-Ldb1), Isl1LBD, and Isl2LBD play a critical role in enabling the formation

---

*This work was supported in part by a grant from the Australian Research Council.

The atomic coordinates and structure factors (code 4JCJ) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 Supported by Australian Postgraduate Awards.

2 Present address: Medical Research Council Laboratory of Molecular Biology, Francis Crick Ave., Cambridge Biomedical Campus, Cambridge CB2 0QH, UK.

3 Present address: Novartis Insts. for BioMedical Research, Klybeckstr. 141, 4057 Basel, Switzerland.

4 Supported by a Senior Research Fellowship from the National Health and Medical Research Council of Australia. To whom correspondence should be addressed. Tel.: 61-2-9351-6025; Fax: 61-2-9351-5858; E-mail: jacqui.matthews@sydney.edu.au.

5 The abbreviations used are: HD, homeodomain; Isl1, Islet 1; Lhx3, LIM homeobox 3; LBD, Lhx3-binding domain; Ldb1, LIM-domain-binding protein 1; LID, LIM interaction domain; LMO, LIM-only; X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; 3-AT, 3-amino-1,2,4-triazole; r.m.s.d., root mean square deviation; TROSY, transverse relaxation optimized spectroscopy; HSQC, heteronuclear single quantum coherence; DBD, DNA-binding domain; AD, activation domain.
of ternary Ldb1-Isl1/2-Lhx3/4 complexes, which in turn up-regulate the expression of motor neuron-specific genes such as Hb9 (see Fig. 1A and Refs. 15 and 21–23).

Ldb1_LBD and Isl1_LBD bind the LIM domains of Lhx3/4 in a similar manner (15, 17), opening up the possibility that Isl1_LBD could mediate interactions with LIM domains from other LIM-HD proteins. Isl1 is critical to the development of motor neurons, pituitary, pancreas, heart, and sensory neurons of the retina and inner ear (18, 24–29). Understanding the roles of this protein in development and the potential ways in which it may be modulated requires a complete appreciation of the behavior of Isl1. This includes the possibility that Isl1 may form an intramolecular interaction through the LBD with its own LIM domains (see Fig. 1B).

The presence of an intramolecular interaction within Isl1 was originally suggested by Sánchez-García et al. (30) based on evidence that the LIM domains in Isl1 could disrupt the DNA binding activity of the protein in vitro. Isl1 showed higher affinity and increased specificity of DNA binding when the LIM domains were removed or denatured. These observations were believed to be the result of an intramolecular interaction between Isl1_LIM1-2 and Isl1_HD. The LIM domains of Lhx3, Mec3 (a Caenorhabditis elegans Lhx1/5 parologue), and Xlim1 (Xenopus Lhx1) have also been observed to negatively regulate DNA binding of their respective homeodomains (12, 31–33). Dawid et al. (34, 35) speculated that the LIM domains of LIM-HD proteins may bind at or near the homeodomain to inhibit DNA binding with inhibition being relieved by interaction of the LIM domains with another binding partner.

This work extends our previous discovery of the Isl1_LBD and characterization of its ability to bind to LIM domains from Lhx3 and Lhx4 (15) through the investigation of binding specificity and modeling of an intramolecular interaction between Isl1_LIM1-2 and Isl1_LBD. We present binding and NMR evidence to demonstrate a weak but specific intramolecular interaction between those domains in Isl1 that provides a possible mechanism by which the LIM domains of Isl1 can influence DNA binding. We have further determined the crystal structure of Isl1_LIM1-2 bound to Ldb1_LBD and use this structure to generate a homology model of an Isl1_LIM1-2-Isl1_LBD complex.

### EXPERIMENTAL PROCEDURES

**Cloning and Protein Expression**—Protein residue numbering refers to the following National Center for Biotechnology Information (NCBI) mouse protein entries: Isl1, NP_067434.3; Ldb1, NP_034827.1; and Lhx3, NP_001034742.1. For LIM1-2 constructs in Fig. 2, the following additional mouse proteins were used, with the residue range for each protein indicated in subscript: Lhx1_W123NP_032524.1; Lhx2_247-174NP_034840.1; Lhx3_28-153NP_034842.2; Lhx5_1-124NP_032525.1; Lhx6_221-224NP_032526.2; Lhx8_209-216NP_034843.2; Lhx9_65-192NP_001036042.1; Isl1_11-138NP_081673.1; Lmx1A_99-154NP_387501.1; and Lmx1B_152NP_034855.1. Genes were cloned from mouse brain tissue or generated as synthetic genes (GenScript USA Inc.). Isl1_LIM1-2 refers to residues 11–138, Isl1_LBD refers to residues 262–301, Ldb1_LID refers to residues 300–330, and Lhx3_LIM1+2 refers to residues 28–153. All other constructs are defined as they appear in the text. Constructs were generated by PCR and sequenced to confirm identity (Australian Genome Research Facility, Westmead Millennium Institute, Sydney, Australia).

Proteins were expressed with a GST tag using pGEX-2T (GE Healthcare) in Escherichia coli BL21 (DE3) cells at 20 °C for 16–20 h. All proteins were purified by glutathione-Sepharose 4B resin (GE Healthcare) and eluted by proteolytic cleavage of the tag with thrombin (Sigma-Aldrich) or the addition of 20 mM reduced glutathione (Sigma-Aldrich). The proteins were additionally purified by size-exclusion chromatography using a Superdex 200 16/60 column (GE Healthcare) in 20 mM Tris (pH 8.5), 150 mM NaCl, and 1 mM dithiothreitol (DTT).

**Yeast Two-hybrid Analysis**—Experiments were carried out using inserts cloned into pGAD10 and modified pGBT9 plasmids, which were co-transformed into Saccharomyces cerevisiae AH109 cells (Clontech), as described previously (13). Successfully co-transformed cells were selected by growth on media lacking leucine and tryptophan. To detect interactions, the following selection conditions were used (all media lacked leucine and tryptophan): low stringency (−His + X-α-Gal), lacking histidine and supplemented with 40 μg/ml 5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside (X-α-Gal); moderate stringency (−His + X-α-Gal + 3-AT), same as for low stringency but supplemented with 1 mM 3-amino-1,2,4-triazole (3-AT); high stringency (−His−Ade), lacking both histidine and adenine. Yeast cultures were normalized to Aν00 = 0.2 and deposited in 2-μl drops at the dilutions indicated.

For yeast mating experiments, AH109 cells were transformed with pGBT9 constructs, Y187 cells were transformed with pGAD10 constructs, and cells were grown overnight at 30 °C in liquid media lacking tryptophan or leucine, respectively. Cultures of AH109 and Y187 cells at Aν00 = 1.0 were spotted on top of each other in an array on rich media and incubated at 30 °C for ~16 h. Spots containing mated diploid cells were replica-plated onto moderate stringency media and incubated at 30 °C for ~48 h.

**Multi-angle Laser Light Scattering**—Purified proteins were subjected to size exclusion chromatography using a Superose 12 10/30 size exclusion column (GE Healthcare) in line with a Wyatt refractometer and MiniDawn multiangle laser light scattering detector (Wyatt Technology). The concentration of protein as determined by the change of refractive index and the intensity of scattered light were used to experimentally determine the molecular mass passing through the detector.

**Thermal Denaturation**—The thermal stability of proteins was determined using the ThermoFluor method (36, 37). Meltin experiments were performed in a 96-well plate format using a 7500 Fast Real-Time PCR System (Applied Biosystems) in tetramethylrhodamine filter mode at a heating rate of 45 s °C−1. Protein was used at a concentration of ~0.33 mg ml−1 in 20 mM Tris (pH 8.5), 150 mM NaCl, and 1 mM DTT and spiked with SYPRO® Orange (Invitrogen) at a final concentration of ~16× (based on the “5000× stock solution” as supplied by Invitrogen). Fluorescence readings were taken at 1° intervals from 25...
An Intramolecular LIM-Peptide Interaction in Isl1

to 95 °C. Data were normalized for basal fluorescence and expressed as a percentage of maximum fluorescence with minimum fluorescence treated as zero.

Nuclear Magnetic Resonance (NMR)—Spectra were acquired at 298 K on a 600- or 800-MHz spectrometer (Bruker Avance III) equipped with a 5-mm TCI CryoProbe (Bruker) as indicated in the figure legends. 15N transverse relaxation optimized spectroscopy-heteronuclear single quantum coherence (TROSY-HSQC) experiments were performed using the standard pulse sequence troseyB3gpsi from the Bruker library. Spectra were processed with TopSpin (Bruker) and analyzed with SPARKY.7

X-ray Structure Determination and Refinement—The crystallization, collection, and processing of data for Isl1LIM1+2−Ldb1LID were described previously (38). The AutoSol (39) applet in the PHENIX software compilation (40) was used to generate experimental phases. The program PHASER (41) was used to identify the zinc atom positions and to generate initial phases with the anomalous data to 3.1-Å resolution. The program RESOLVE (42) was used for density modification and subsequent map and subsequently revised using Coot (43).

For refinement of the model, the data were reprocessed and scaled using HKL-2000 (44) to include only the first 360 images of 0.5° oscillation each, which corresponds to 180° total oscillation and half of the total collected images. Refinement using the new data set was carried out using REFMAC (45) with isotropic temperature factors, TLS groups, and metal-ligand bond restraints between the zinc atoms and coordinating atoms (46). The MOLPROBITY server was used for structure validation, and through decreased entropy loss brought about by the domains being part of the same polypeptide chain. Before investigating the existence of this weak intramolecular interaction, we tested whether any intramolecular interaction could be physically accommodated in the native protein (in which the homeodomain lies between LIM domains and LBD; Fig. 1B). Thus, we generated a construct of Isl1 that encompassed the LIM domains, homeodomain, and LBD (Isl111–291) and a chimera in which the Isl1LBD was substituted by Ldb1LID (Isl1−291LBD) and a chimeric in which the Isl1LBD was substituted by Ldb1LID (Isl1+Ldb1; Fig. 2A). So despite a lack of evidence for an intramolecular interaction between the LIM domains of Isl1 (Isl1LIM1+2) and Isl1LBD (Fig. 2A and Ref. 15), we hypothesized that Isl1LIM1+2 might be able to bind Isl1LIM1+2 in an intramolecular fashion. That is, an intrinsically weak interaction between those domains would be enhanced through decreased entropy loss brought about by the domains being part of the same polypeptide chain. Before investigating the existence of this weak intramolecular interaction, we tested whether any intramolecular interaction could be physically accommodated in the native protein (in which the homeodomain lies between LIM domains and LBD; Fig. 1B). Thus, we generated a construct of Isl1 that encompassed the LIM domains, homeodomain, and LBD (Isl111–291) and a chimera in which the Isl1LBD was substituted by Ldb1LID (Isl1−291LBD) and a chimeric in which the Isl1LBD was substituted by Ldb1LID (Isl1+Ldb1; Fig. 2, B and C). This differs from “tethered complexes” (Fig. 2D) characterized in previous studies in which the LIM domains from LIM-HD/LMO proteins are tethered to Ldb1LID or Isl1/2LBD via a flexible linker (e.g. Refs. 13, 15, 17, 53, and 54). We used the yeast two-hybrid system to compare the binding of these constructs to known protein partners. Lhx3LIM1+2 (which can bind Isl1LBD or Ldb1LID) and Ldb1LID (which can bind Isl1LIM1+2) (15). Whereas Isl111−291 was able to bind robustly to each of Lhx3LIM1+2 and Ldb1LID, under all selection conditions, Isl1+Ldb1 did not bind its targets under any conditions (Fig. 2E). This result is consistent with a strong intramolecular interaction between Isl1LIM1+2 and Ldb1LID in the chimera that prevents intermolecular interaction with either Lhx3LIM1+2 or Ldb1LID. That is, an LBD-like sequence can physically contact Isl1LIM1+2 despite the presence of the intervening homeodomain in the primary sequence, suggesting that an intramolecular LIM-LBD interaction could physically occur in the Isl1 protein. However, the data for Isl111−291 indicate that the competing intermolecular interactions Isl1LIM1+2-Lhx3LIM1+2 and Isl1LIM1+2-Ldb1LID would be much stronger than a putative intramolecular interaction.

7 T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco.
An Intramolecular LIM-Peptide Interaction in Isl1

**Generation of Tethered Complexes and Negative Control LBDs**—Stable forms of LIM-HD and LMO proteins are generated by tethering the LIM domains to an interaction partner such as Ldb1LID or Isl1LBD (15, 51, 52, 55, 56) (Fig. 2D). This approach works well because those peptide-like structures bind their target LIM domains in a head-to-tail manner (i.e. the C terminus of the LID/LBD lies near the N terminus of the LIM domains and vice versa). We engineered Isl1LIM1+2 and Ldb1LID-Isl1LBD constructs. We generated Isl1LIM1+2-Lbd1LID and Isl1LBD-Isl1LIM1+2 constructs. Changing the order of binding partners in this type of complex can affect the apparent stability (57), so throughout this study we compare only like with like for tethered complexes (e.g. LIMs first with LIMs first or LIMs second with LIMs second) as indicated in the figure legends. For simplicity, we henceforth refer to the tethered complexes of Isl1LIM1+2 with Ldb1LID and Isl1LBD as Isl1-Ldb1LID and Isl1-Isl1LBD, respectively, regardless of the order of binding partners.

To establish whether Isl1LBD binds Isl1LIM1+2 in a specific or nonspecific manner, we generated two negative control LBDs based on Isl1LBD. In “Switched,” the positions of the LIM1- and LIM2-binding motifs within the LBD sequence were switched (Fig. 3A). As the two binding motifs from LID/LBD-like peptides are thought to not be interchangeable (58), this design should prevent Switched from contacting both LIM1 and LIM2 simultaneously and thereby reduce binding affinity. Note that for this construct an interaction could be mediated with LIM1 or LIM2. “Scrambled” was generated by randomizing the order of the 30 residues of Isl1LBD. This design maintains the same overall physical properties but should disrupt any contacts that confer specific binding. Switched and Scrambled were tested for binding against Lhx3LIM1+2 in a yeast two-hybrid assay, and as expected, both showed reduced binding activity (Fig. 3B). Switched and Scrambled were incorporated into tethered Isl1LIM1+2-LBD constructs as described above, resulting in Isl1-Switched and Isl1-Scrambled.

**Blocking Ldb1LID with Tethered Complexes**—We used a competitive yeast two-hybrid approach to establish the presence of a weak intramolecular interaction in Isl1. The assay used Isl1LBD constructs (including wild-type, switched, and scrambled versions of LBD) fused to the Gal4 activation domain (AD) and Ldb1LID peptides (including a series of Ldb1LID peptides that have reduced affinity for Isl1LIM1+2, Ref. 15) fused to the Gal4 DNA-binding domain (DBD). Ldb1LID Peptides that can effectively compete with Isl1LBD for binding to Isl1LIM1+2 will.

---

**FIGURE 2.** Identifying new interactions for Isl1LBD. **A**, yeast two-hybrid data for Ldb1LID and Isl1LBD interactions with the LIM domains (LIM1 + 2) of the 12 mammalian LIM-HD proteins as indicated. Mated AH109/Y187 yeast cells transformed with pGBT9/pGAD10 vectors, which encode the Gal4 DBD and AD, respectively. Co-transformed cells were tested for growth under moderate stringency conditions (−His + X-α-Gal + 3-AT). Data are shown for pGBT9-LIMHDLIM1+2 and pGAD10-Ldb1LID/Isl1LBD pairs. pGBT9-only (DBD-only) and pGAD10-only (AD-only) negative controls contain no insert in those vectors. **B**, schematics of Isl1, Ldb1, and chimeric constructs. LIM1 and LIM2, first and second LIM domain, respectively. **C**, hypothesized domain arrangement of Isl111–291 and Isl1 + Ldb1 constructs. **D**, schematic of tethered Isl111–291Ldb1LID and Isl1LIM1+2-Isl1LBD complexes. E, yeast two-hybrid data testing binding of Isl1, Ldb1, and chimeric constructs with Ldb1 and Lhx3. AH109 yeast cells co-transformed with pGBT9/pGAD10 vectors as indicated were tested for growth under moderate (−His (−H) + X-α-Gal + 3-AT) and stringent (−His-Ade (−H−A)) selection conditions and the indicated starting densities of yeast cells (A600 nm).

---
disrupt the intramolecular interaction and result in activation of reporter genes, whereas peptides that have a reduced affinity for \( \text{Is}l1_{\text{LIM1+2}} \) and cannot effectively compete with \( \text{Is}l1_{\text{LBD}} \) for binding to \( \text{Is}l1_{\text{LIM1+2}} \) will not activate reporter genes (Fig. 3C). Thus, in terms of assessing the relative strength of the intramolecular interaction, reduced levels of yeast growth indicate a stronger interaction, and robust levels of growth indicate a weaker interaction.

Wild-type \( \text{Ldb1}_{\text{LID}} \) was able to bind all three tethered constructs under both moderate and high stringency conditions,
indicating that neither Isl1LBD nor the negative control LBDs can block the interaction of Ldb1LID with Isl1LIM1+2 (Fig. 3D). Similarly, T323A did not significantly disrupt binding to any of the tethered complexes tested, reflecting its high affinity for Isl1LIM1+2, as observed previously (15). Conversely, V303A/I322A abolished binding to the tethered constructs under all conditions. All other mutants of Ldb1LID bound Isl1-Isl1LBD very weakly or not at all, indicating that an intramolecular Isl1-Isl1LBD interaction exists that can out-compete weaker mutant Ldb1LID/Isl1LIM1+2 interactions. In contrast, the mutant peptides showed much higher levels of binding to Isl1-Scrambled and Isl1-Switched with the latter showing the highest levels of binding. These data indicate that the relative strengths of the binding affinities of Isl1LBD for Isl1LIM1+2 are wild type > scrambled > switched.

Production and Assessment of Isolated Tethered Complexes— Recombinant Isl1-Isl1LBD and Isl1-Ldb1LID were expressed in E. coli and purified albeit at much lower quantities for Isl1-Isl1LBD (~0.3 mg of protein/liter of culture) compared with Isl1-Ldb1LID (3–4 mg; Ref. 59). In contrast, Isl1-Switched and Isl1-Scrambled either formed large aggregates or were degraded, precluding biophysical or structural analysis as isolated tethered complexes. Isl1-Ldb1LID and Isl1-Isl1LBD were each determined to be largely monomeric (~20 kDa) by multivite laser light scattering; however, small amounts of higher molecular weight species (25–35 kDa) were detected for Isl1-Isl1LBD (Fig. 4A). Based on thermal denaturation, the Isl1-Isl1LBD complex was much less stable than Isl1-Ldb1LID (Fig. 4B). Both complexes yielded denaturation curves consistent with cooperative unfolding but with different midpoints of unfolding, ~65 and 85 °C, for Isl1-Isl1LBD and Isl1-Ldb1LID, respectively.

The Specificity of the Intramolecular Interaction—Fortunately, it was possible to express and purify modest amounts of 15N-labeled GST-tagged Isl1-Switched and Isl1-Scrambled. Thus, we recorded 15N-1H HSQC NMR spectra of the GST-tagged Isl1-peptide complexes to assess the folded states of these constructs. This approach has been used successfully in the past to assess the stability and folded state of GST-tagged proteins (60, 61). In those studies, the linker between the GST and the protein of interest was flexible, and the proteins of interest had a low molecular mass (~10 kDa or smaller). Traditional HSQC spectra are very poor at detecting signals from species with a high molecular mass due to slower tumbling of molecules in solution, so very few peaks were observed for the GST component (at ~50 kDa for dimeric GST) of those constructs, and the spectra predominantly reported narrow, sharp peaks for the faster tumbling small protein of interest. In comparison, the various tethered complexes containing Isl1LIM1+2 are larger (~20 kDa) and expected to be rod-shaped. We collected 15N-1H TROSY-HSQC data (Fig. 5) because traditional HSQC spectra recorded on our GST fusion proteins displayed very few peaks (<5% of the number of peaks observed by the TROSY experiments). TROSY experiments exploit a partial cancellation of the dipole-dipole coupling of the amide group and chemical shift anisotropy of the nitrogen atoms for certain interactions (62). This increases signal resolution and allows the measurement of larger molecular mass proteins.

The spectra of the GST-tagged tethered complexes (GST-Isl1-Ldb1LID, GST-Isl1-Isl1LBD, GST-Isl1-Switched, and GST-Isl1-Scrambled) were compared with that of GST alone (Fig. 5). The 15N-1H TROSY-HSQC spectrum of GST yielded many more peaks than the standard 15N-1H HSQC (60), but these peaks are predominantly clustered in a few areas of the spectrum, and all four GST-tagged tethered complexes contained additional peaks when compared with GST alone. The spectrum of GST-Isl1-Ldb1LID showed a reasonable number (~75) of well dispersed and sharp non-GST peaks indicative of a highly folded protein (Fig. 5A). The spectrum of GST-Isl1-Isl1LBD showed many of the same peaks as that of GST-Isl1-Ldb1LID (~50), but these peaks were generally broader (Fig. 5B). In contrast, the spectra for GST-Isl1-Switched and GST-Isl1-Scrambled contained far fewer peaks. Other than those for tryptophan side chains (15N, ~130 ppm; 1H, ~10 ppm), the only non-GST peaks that are visible for these proteins are poorly dispersed and lie in the “unfolded” regions of the spectra (Fig. 5, C and D) (63), suggesting that these constructs are not well folded and are probably molten globule-like. On the whole,
these data indicate that Ldb1LID and Isl1LBD can form specific interactions with the LIM domains of Isl1, but the negative control LBDs, Switched and Scrambled, do not.

The Crystal Structure of Isl1-Ldb1LID—We attempted to determine the three-dimensional structure of Isl1-Ldb1LID and Isl1-Isl1LBD to better understand the nature of the intermolecular and intramolecular Isl1-peptide interactions. Isl1-Isl1LBD did not yield diffraction quality crystals, but Isl1-Ldb1LID was successfully crystallized, and diffraction data were collected as described previously (38).

Experimental phases were generated from this diffraction data by single wavelength anomalous dispersion at the zinc absorption edge (1.2186 Å) to a resolution of 3.1 Å (Table 1). To improve the quality of the data and reduce potential errors introduced by radiation damage, we reprocessed it to take the first 360 frames (instead of 720) and extended the resolution to 3.0 Å (Table 1). These data were used for refinement of the Isl1-Ldb1LID model. The \( R \) work and \( R \) free of the final model correspond well with the resolution of the data and the amount of the electron density that can be adequately modeled (~20% of total mass is missing; Table 1; Protein Data Bank code 4JCJ).

Three molecules of Isl1-Ldb1LID exist in the asymmetric unit and exhibit the same basic structure. The LIM domains of Isl1 in this complex are very similar to those of other LIM-only and LIM-HD complexes (13, 15, 17, 54). The two LIM domains each contain two zinc-coordinating modules (i.e. coordinated atoms Zn1–4), which comprise two \( \beta \)-hairpins followed by a short \( \alpha \)-helix (Fig. 6A, blue ribbon). The LIM domains are separated by a short region known as the “hinge.” Ldb1LID binds both LIM domains in an extended, head-to-tail fashion but does so in a bipartite manner in which two peptide segments of eight and 10 residues contact the two separate LIM domains (green ribbon). Residues Met302–Leu309 contact Isl1LIM2 and residues Asp318–Asn327 contact Isl1LIM1. The intervening residues, many of which are absent from the electron density, correspond to the “spacer,” which if present would sit across from the hinge of Isl1. The LIM1-binding region forms a single extended \( \beta \)-strand that packs in an antiparallel manner against the second \( \beta \)-hairpin of the first and second zinc-coordinating modules (which contain Zn1 and Zn2, respectively). These \( \beta \)-sheet structures overlap such that Ldb1318–322 contacts Isl156–60 and Ldb1321–327 contacts Isl127–33. In contrast, the LIM2-binding region forms a \( \beta \)-strand-like structure with the second \( \beta \)-hairpin of the third zinc-coordinating module (Zn3) and a \( \beta \)-strand with the second \( \beta \)-hairpin of the fourth (Zn4). Ldb1308 contacts Isl192 and Ldb1302–304 contacts residues of Isl1117–120 (green ribbon), resulting in a single, three-stranded \( \beta \)-sheet being created at the fourth zinc-coordinating module. The buried surface area of the interface between Isl1LIM2 and Ldb1LID is \(~1290 \text{ Å}^2\).

The three molecules in the asymmetric unit (chains A, B, and C; chain B is shown in Fig. 6A) are largely elongated, exhibiting angles at the hinge/spacer of 169° for A, 175° for B, and 176° for C. Chains B and C are the most similar (root mean square deviation (r.m.s.d.), 0.61 Å), whereas chain A is more divergent from B and C, having r.m.s.d. values of 1.45 and 1.67 Å, respectively.
An Intramolecular LIM-Peptide Interaction in Isl1

When the individual LIM domains or the corresponding LIM-binding domains of the LBDs are aligned, the molecules are much more similar (Fig. 6, B and C). In general, the LIM1 halves of the molecules (r.m.s.d. values, 0.31–0.35 Å) are more similar than the LIM2 halves (r.m.s.d. values, 0.42–0.55 Å).

Of the 173 residues in Isl1-Ldb1<sub>LID</sub>, 23 are not found in any of the chains, whereas an additional five are only found in chains B and C, one is in chains A and C, and another one is only in chain C. The missing residues from Isl1-Ldb1<sub>LID</sub> are largely from the linker and termini of the protein but are also residues of the spacer of Ldb1<sub>LID</sub> not adjacent to the LIM-binding regions (311–315). Of those residues that were modeled, ~25% are missing side chains (for details of missing residues and atoms, see Protein Data Bank code 4JC).

One of the crystal packing contacts is mediated by a disulfide bond between Cys-56 residues from pairs of Isl1 molecules (Fig. 6D). All protein samples used for crystallization contained DTT; however, samples appeared to crystallize more readily with DTT; however, samples appeared to crystallize more readily with DTT, but in a poor way from the Isl1-LIM1-Ldb1<sub>LID</sub> structure to the corresponding binding residues from Isl1<sub>LBD</sub> using the Lhx3-Isl1-Ldb1<sub>LID</sub> structures as a guide. Based on this homology model, Isl1<sub>LBD</sub>, Isl1-LIM1<sub>LID</sub> is predicted to have a surface area (~1330 Å²) similar to that of Ldb1<sub>LID</sub>-Isl1-LIM1<sub>LID</sub> (~1290 Å²). The model of Isl1-Isl1<sub>LBD</sub> reveals no apparent impediments to binding (Fig. 6, F and G), but several differences suggest an overall poorer complementarity between Isl1-LIM1<sub>LID</sub> and Isl1-LDB<sub>LID</sub>, resulting in the weaker intramolecular interaction. Isl1-LDB<sub>LID</sub>-Isl1-LIM1<sub>LID</sub> forms 19 interdomain hydrogen bonds compared with 26 for Ldb1<sub>LID</sub>-Isl1-LIM1<sub>LID</sub> and the residues of Ldb1<sub>LID</sub> that are buried in grooves between the zinc-binding modules of each LIM domain are smaller in Isl1-LBD (Ldb1<sub>Val-304</sub> corresponds to Isl1<sub>Ala-267</sub>, and Ldb1<sub>Ile-322</sub> corresponds to Isl1<sub>Val-282</sub>).

DISCUSSION

The structure of Isl1-Ldb1<sub>LID</sub> can be compared with existing structures of Lhx3-Ldb1<sub>LID</sub>, Lmo2-Ldb1<sub>LID</sub>, Lmo4-Ldb1<sub>LID</sub>, Lhx3-Isl1-Ldb1<sub>LID</sub>, Lhx4-Isl2-Ldb1<sub>LID</sub>, and Lhx4-Isl4-Ldb1<sub>LID</sub> complexes (13, 15, 17, 54). The tethered complexes generally crystallize in different conformations as a result of flexibility in the hinge/spacer region (17, 54, 59), so it is only meaningful to compare the structures of individual LIM-LID/LBD “half-complexes,” the folds of which are highly conserved (Table 2). In this context, the complexes are very similar, although the lengths of well defined β-strands in Ldb1 differ (Fig. 7A). In the LIM1 halves of the complexes, Ldb1 forms one long β-strand with Isl1, two shorter β-strands with Lmo2 and Lmo4, and a single short β-strand with Lhx3 (Fig. 7A). For the LIM2 halves of the complexes, Ldb1<sub>LID</sub> forms a single short β-strand with Isl1 and Lhx3 and two short β-strands with Lmo2 and Lmo4 structures (Fig. 7A).

Isl1, Lmo2, and Lmo4 all bind Ldb1 preferentially through their LIM1 domains, whereas Lhx3 preferentially binds through its LIM2 domain (14–17). A possible contribution for this difference can be seen around the key LIM1-binding residue, Ldb1<sub>ile-322</sub>. Whereas LIM1 in Isl1, Lmo2, and Lmo4 provides the perfect binding surface for the mt rotamer of Ldb1<sub>ile-322</sub>, the equivalent surface in Lhx3<sub>LIM1</sub> binds Ldb1<sub>ile-322</sub> as the less common pt rotamer (Fig. 7, B and C). These differences appear to be derived from Lhx3 having a shorter Loop 7, which connects the first and second β-hairpins of the second zinc-coordinating module and lies adjacent Ldb1<sub>ile-322</sub> (Fig. 7D). The longer Loop 7 also appears to enable the formation of the additional backbone-backbone hydrogen bonds across the Ldb1<sub>LID</sub>-Isl1/Lmo2/Lmo4-LIM1 interfaces that result in increased secondary structure in Ldb1 for Isl1/Isl1/Ldb1 compared with Lhx3 at LIM1 (Fig. 7A).

Our data support the presence of an intramolecular interaction in Isl1 whereby the LBD from this protein can interact in a specific manner with its own LIM domains. Our yeast two-hybrid and thermal denaturation data of the Isl1-Isl1<sub>LBD</sub> tethered complex indicate that the intramolecular interaction is weak compared with intermolecular Isl1-Ldb1 binding, suggesting that the intermolecular interaction is favored.

| TABLE 1  |
| --- |
| X-ray data and model statistics for Isl1-Ldb1<sub>LID</sub> |
| Values for the highest resolution shell are given in parentheses. |

| Phasing statistics | Resolution range | 50.0–3.1 |
| --- | --- | --- |
| FOM<sub>PHASES</sub> (centric/acentric)<sup>a</sup> | 0.43/0.43 |
| FOM<sub>RESOLVE</sub> (centric/acentric) | 0.55/0.51 |

| Scaling statistics for refinement | Wavelength (Å) | 1.2816 |
| --- | --- | --- |
| Space group | P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> |
| Unit cell parameters (Å) | a = 57.3, b = 56.6, c = 179.3 |
| Resolution limits (Å) | 50.0–3.0 (3.07–3.00) |
| Completeness (%) | 100.0 (100.0) |
| Wilson B factor (Å<sup>2</sup>) | 59.6 |
| Observed reflections | 84,886 |
| Unique reflections | 11,622 |
| Redundancy | 6.9 (7.0) |
| R<sub>merge</sub> | 0.160 (0.661) |
| R<sub>free</sub> | 8.5 (2.8) |

| Model refinement | R<sub>work</sub> | 25.4 (35.8) |
| --- | --- | --- |
| R<sub>free</sub> | 29.9 (37.6) |
| No. of reflections used in refinement | 11,622 (800) |
| No. of reflections in the test set | 596 (453) |
| Protein atoms (including zinc) | 3,095 |
| Solvent molecules | 5 |
| r.m.s.d. bond length (Å) | 0.01 |
| r.m.s.d. bond angles (°) | 1.30 |
| Mean protein ϕ β factor (Å<sup>2</sup>) | 46.0 |

| Ramachandran plot, residues in | Favored regions (%) | 91.3 |
| --- | --- | --- |
| Allowed regions (%) | 8.2 |
| Disallowed regions (%) | 0.5 |

<sup>a</sup> FOM (figure of merit) is the expected value (probability-weighted average) of the cosine of the phase error.

<sup>b</sup> R<sub>merge</sub> = Σ<sub><i>i</sub></sub>Σ<sub><i>hkl</sub></sub>[I<sub><i>obs</sub></i> - 1/|I<sub><i>calc</sub></i>|]|I<sub><i>calc</sub></i>]/Σ<sub><i>i</sub></sub>Σ<sub><i>hkl</sub></sub>[I<sub><i>obs</sub></i>].

<sup>c</sup> R<sub>work</sub> = 1/Σ<sub><i>hkl</sub></sub>[F<sub><i>obs</sub></i> - F<sub><i>calc</sub></i>]/Σ<sub><i>hkl</sub></i>[F<sub><i>obs</sub></i>].

<sup>d</sup> R<sub>free</sub> is R<sub>work</sub> for the 5% validation set.
The origin of an intramolecular interaction between Isl1LBD and Isl1LIM1 is unclear. It may simply reflect that Isl1LBD binds the LIM domains of Lhx3 and Lhx4 on the same surface as Ldb1LID, which in turn binds Isl1LIM1. Alternatively, it is possible that intramolecular interactions involving LIM domains were a feature of ancestral LIM-HD proteins in which case the Isl1LBD may have subsequently evolved to mediate intramolecular interactions with Lhx3/4.

Regardless, the intramolecular interaction could play one or more positive roles in regulating the function of Isl1, many of which benefit from having a weak rather than a strong binding affinity. Uncomplexed LIM domains from Isl1 are likely to be “sticky” because of exposed hydrophobic surfaces. Isl1LBD might shield the LIM domains and prevent Isl1LIM1 from nonspecifically binding other proteins in the cell or prevent the protein from being targeted for proteolytic degradation. The intramolecular interaction may as suggested in concept by Dawid (35) prevent Isl1 from binding DNA in the absence of its...
other protein partners (Fig. 8A) but as a weak interaction allow readily for productive binding (Fig. 8B). We suggested previously that preferential binding of Isl1LIM1 to Ldb1 and of Lhx3LIM2 to Ldb1 and Isl1 might facilitate an exchange mechanism to explain how the Ldb1/Lhx3 complex could be disrupted by Isl1 (15). An intramolecular interaction between Isl1LBD and Isl1LIM1 allows for protection of Isl1LBD and Isl1LIM1 in this scenario with both the Isl1-Ldb1 and Isl1-Lhx3 interactions in the Ldb1-Isl1-Lhx3 ternary complex forming in parallel (Fig. 8C).

The possibility of an intramolecular interaction in Isl1 opens up the potential of similar interactions in other LIM-HD proteins especially if the origin of Isl1LBD occurred in an ancestral LIM-HD. The largely uncharacterized and unstructured C-terminal regions of LIM-HD proteins stand in stark contrast to the highly structured and conserved N-terminal regions. It seems probable that in at least some LIM-HD proteins these regions contain a non-conserved, intrinsically unstructured sequence that might very weakly but specifically interact with the LIM domains in an intramolecular fashion.
An Intramolecular LIM-Peptide Interaction in Isl1

Acknowledgments—We thank Soumya Joseph for providing the 15N-labeled GST protein used in the NMR experiments. This research was undertaken on the MX2 beamline at the Australian Synchrotron, Victoria, Australia.

REFERENCES
1. Bachy, I., Vernier, P., and Retaux, S. (2001) The LIM-homeodomain gene family in the developing Xenopus brain: conservation and divergences with the mouse related to the evolution of the forebrain. J. Neurosci. 21, 7620–7629
2. Tsuchida, T., Ensini, M., Morton, S. B., Baldassare, M., Edlund, T., Jessell, T. M., and Pfaff, S. L. (1994) Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. Cell 79, 957–970
3. Bachy, I., Failli, V., and Retaux, S. (2002) A LIM-homeodomain code for development and evolution of forebrain connectivity. Neuronreport 13, A22–A27
4. Lumsden, A. (1995) Neural development. A ‘LIM code’ for motor neuron cell type specification. Nature 374, 483–486
5. Ahlgren, V., Pfaff, S. L., Jessell, T. M., Edlund, T., and Edlund, H. (1997) Independent requirement for Isl1 in formation of pancreatic mesenchyme and islet cells. Nature 385, 257–260
6. Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., Mackay, J. P., Kwan, A. H., Sum, E. Y., Visvader, J. E., and Bhati, M., Lee, C., Nancarrow, A. L., Lee, M., Craig, V. J., Bach, I., Guss, J. M., and Matthews, J. M. (2010) Structural basis for partial redundancy in a rhombotin and LIM homeodomain interacting protein, is expressed in the mouse related to the evolution of the forebrain. J. Neurosci. 27, 12707–12720
7. Ahlgren, V., Pfaff, S. L., Jessell, T. M., Edlund, T., and Edlund, H. (1997) Independent requirement for Isl1 in formation of pancreatic mesenchyme and islet cells. Nature 385, 257–260
8. Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., Mackay, J. P., Kwan, A. H., Sum, E. Y., Visvader, J. E., and Bhati, M., Lee, C., Nancarrow, A. L., Lee, M., Craig, V. J., Bach, I., Guss, J. M., and Matthews, J. M. (2010) Structural basis for partial redundancy in a rhombotin and LIM homeodomain interacting protein, is expressed in the mouse related to the evolution of the forebrain. J. Neurosci. 27, 12707–12720
9. Ahlgren, V., Pfaff, S. L., Jessell, T. M., Edlund, T., and Edlund, H. (1997) Independent requirement for Isl1 in formation of pancreatic mesenchyme and islet cells. Nature 385, 257–260
10. Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., Mackay, J. P., Kwan, A. H., Sum, E. Y., Visvader, J. E., and Bhati, M., Lee, C., Nancarrow, A. L., Lee, M., Craig, V. J., Bach, I., Guss, J. M., and Matthews, J. M. (2010) Structural basis for partial redundancy in a rhombotin and LIM homeodomain interacting protein, is expressed in the mouse related to the evolution of the forebrain. J. Neurosci. 27, 12707–12720
11. Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., Mackay, J. P., Kwan, A. H., Sum, E. Y., Visvader, J. E., and Bhati, M., Lee, C., Nancarrow, A. L., Lee, M., Craig, V. J., Bach, I., Guss, J. M., and Matthews, J. M. (2010) Structural basis for partial redundancy in a rhombotin and LIM homeodomain interacting protein, is expressed in the mouse related to the evolution of the forebrain. J. Neurosci. 27, 12707–12720
12. Deane, J. E., Ryan, D. P., Sunde, M., Maher, M. J., Guss, J. M., Visvader, J. E., Mackay, J. P., Kwan, A. H., Sum, E. Y., Visvader, J. E., and Bhati, M., Lee, C., Nancarrow, A. L., Lee, M., Craig, V. J., Bach, I., Guss, J. M., and Matthews, J. M. (2010) Structural basis for partial redundancy in a rhombotin and LIM homeodomain interacting protein, is expressed in the mouse related to the evolution of the forebrain. J. Neurosci. 27, 12707–12720
40. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Saccaccini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954

41. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2001) Phaser crystallographic software. J. Appl. Crystallogr. 34, 776–782

42. Terwilliger, T. C. (2000) Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. 56, 965–972

43. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132

44. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326

45. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 53, 240–255

46. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. D Biol. Crystallogr. 57, 122–133

47. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Co geometry: ϕ, ψ and Cβ deviation. Proteins 50, 437–450

48. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797

49. Wallace, A. C., Wootton, J. C., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 8, 127–134

50. Bhat, M., Lee, M., Nancarrow, A. L., Bach, I., Guss, J. M., and Matthews, J. M. (2008) Crystallization of an Lhx3-Is1 complex. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 64, 297–299

51. Gadd, M. S., Langley, D. B., Guss, J. M., and Matthews, J. M. (2009) Crystallization and diffraction of an Lhx3-Is1 complex. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 65, 151–153

52. Lee, C., Nancarrow, A. L., Bach, I., Mackay, J. P., and Matthews, J. M. (2005) 1H, 15N and 13C assignments of an intramolecular Lhx3-Ldb1 complex. J. Biomol. NMR 33, 198

53. Ryan, D. P., Sund, M., Kwan, A. H., Marianayagam, N. I., Nancarrow, A. L., Vanden Hoven, R. N., Thompson, L. S., Baca, M., Mackay, J. P., Visvader, J. E., and Matthews, J. M. (2006) Identification of the key LMO2-binding determinants on Ldb1. J. Mol. Biol. 359, 66–75

54. El Omari, K., Hoosdally, S. J., Tuladhar, K., Karia, D., Vyas, P., Patient, R., Porcher, C., and Mancini, E. J. (2011) Structure of the leukemia oncogene LMO2: implications for the assembly of a hematopoietic transcription factor complex. Blood 117, 2146–2156

55. Deane, J. E., Maher, M. J., Langley, D. B., Graham, S. C., Visvader, J. E., Guss, J. M., and Matthews, J. M. (2003) Crystallization of FLINC4, an intramolecular LMO4-LDB1 complex. Acta Crystallogr. D Biol. Crystallogr. 59, 1484–1486

56. Deane, J. E., Sum, E., Mackay, J. P., Lindeman, G. J., Visvader, J. E., and Matthews, J. M. (2001) Design, production and characterization of FLIN2 and FLIN4: the engineering of intramolecular LDB1-LMO complexes. Protein Eng. 14, 493–499

57. Jeffries, C. M., Graham, S. C., Stokes, P. H., Collyer, C. A., Guss, J. M., and Matthews, J. M. (2006) Stabilization of a binary protein complex by intein-mediated cyclization. Protein Sci. 15, 2612–2618

58. Matthews, J. M., and Potts, J. R. (2013) The tandem β-zipper: modular binding of tandem domains and linear motifs. FEBS Lett. 587, 1164–1171

59. Bhati, M., Lee, C., Gadd, M. S., Jeffries, C. M., Kwan, A., Whitten, A. E., Trewella, J., Mackay, J. P., and Matthews, J. M. (2012) Solution Structure of the LIM-homeodomain transcription factor complex Lhx3/Ldb1 and the effects of a pituitary mutation on key Lhx3 interactions. PLoS One 7, e40719

60. Liew, C. K., Gamsjaeger, R., Mansfield, R. E., and Mackay, J. P. (2008) NMR spectroscopy as a tool for the rapid assessment of the conformation of GST-fusion proteins. Protein Sci. 17, 1630–1635

61. Wissmuller, S., Font, J., Liew, C. W., Cram, E., Schroeder, T., Turner, J., Crossley, M., Mackay, J. P., and Matthews, J. M. (2011) Protein-protein interactions: analysis of a false positive GST pulldown result. Proteins 79, 2365–2371

62. Pervushin, K., Riek, R., Wider, G., and Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl. Acad. Sci. U.S.A. 94, 12366–12371

63. Kwan, A. H., Mobli, M., Gooley, P. R., King, G. F., and Mackay, J. P. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist. FEBS J. 278, 687–703

64. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. I., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948