Solution Structure of the LIM-Homeodomain Transcription Factor Complex Lhx3/Ldb1 and the Effects of a Pituitary Mutation on Key Lhx3 Interactions

Mugdha Bhati ña, Christopher Lee ñ, Morgan S. Gadd ñ, Cy M. Jeffries ñb, Ann Kwan, Andrew E. Whitten ñc, Jill Trewhella, Joel P. Mackay, Jacqueline M. Matthews*

School of Molecular Bioscience, University of Sydney, Sydney, New South Wales, Australia

Abstract

Lhx3 is a LIM-homeodomain (LIM-HD) transcription factor that regulates neural cell subtype specification and pituitary development in vertebrates, and mutations in this protein cause combined pituitary hormone deficiency syndrome (CPHDS). The recently published structures of Lhx3 in complex with each of two key protein partners, Isl1 and Ldb1, provide an opportunity to understand the effects of mutations and posttranslational modifications on key protein-protein interactions. Here, we use small-angle X-ray scattering of an Ldb1-Lhx3 complex to confirm that in solution the protein is well represented by our previously determined NMR structure as an ensemble of conformers each comprising two well-defined halves (each made up of LIM domain from Lhx3 and the corresponding binding motif in Ldb1) with some flexibility between the two halves. NMR analysis of an Lhx3 mutant that causes CPHDS, Lhx3(Y114C), shows that the mutation does not alter the zinc-ligation properties of Lhx3, but appears to cause a structural rearrangement of the hydrophobic core of the LIM2 domain of Lhx3 that destabilises the domain and/or reduces the affinity of Lhx3 for both Ldb1 and Isl1. Thus the mutation would affect the formation of Lhx3-containing transcription factor complexes, particularly in the pituitary gland where these complexes are required for the production of multiple pituitary cell types and hormones.

Introduction

Lhx3 (LIM homeobox protein 3) is essential for specification of many pituitary and neural cell types [1,2,3]. Humans that carry mutations in Lhx3 present with combined pituitary hormone deficiency syndrome (CPHDS) [4,5,6,7,8]. Depending on the site of mutation, affected patients can also exhibit hearing loss and skeletal malformations of the upper body [8,9]. Lhx3 is from the LIM-homeodomain transcription factor family, members of which contain a pair of closely spaced N-terminal LIM domains followed by a central homeodomain (Figure 1A) [23,24,25]. These tethered complexes are hereafter referred to as Ldb1-Lhx3 and Lhx3-Isl1 (the order of the names indicates the order of the domains in the complex). Our recently determined structures of Ldb1-Lhx3 and Lhx3-Isl1 show that Ldb1 LBD and Isl1 LBD interact with Lhx3 in an essentially identical manner [13], with the two binding partners forming extended chains and contacting the same sites across both LIM domains of binding protein [17]. The DNA-binding specificity of Lhx3 is context dependent, and varies according to the protein isoform [18,19], or whether Lhx3 is in acting in concert with a protein partner [2,20,21,22]. Disease-causing mutations or post-translational modifications, including phosphorylation, of the LIM domains are likely to affect the biological activity of Lhx3 by modulating protein-protein interactions and modulating binding to DNA targets [5,6,10].

The isolated LIM domains from Lhx3 (Lhx3 LIM1+2) tend to be insoluble and/or aggregate, but soluble stable “tethered complexes” can be engineered in which the LIM interaction domain of Ldb1 (Ldb1 LID), or the Lhx3-binding domain from Isl1 (Isl1 LBD) are fused to Lhx3 LIM1+2 via a flexible glycine-serine linker (Figure 1B) [23,24,25]. These tethered complexes are hereafter referred to as Ldb1-Lhx3 and Lhx3-Isl1 (the order of the names indicates the order of the domains in the complex). Our recently determined structures of Ldb1-Lhx3 and Lhx3-Isl1 show that Ldb1 LBD and Isl1 LBD interact with Lhx3 in an essentially identical manner [13], with the two binding partners forming extended chains and contacting the same sites across both LIM domains of binding protein [17]. The DNA-binding specificity of Lhx3 is context dependent, and varies according to the protein isoform [18,19], or whether Lhx3 is in acting in concert with a protein partner [2,20,21,22]. Disease-causing mutations or post-translational modifications, including phosphorylation, of the LIM domains are likely to affect the biological activity of Lhx3 by modulating protein-protein interactions and modulating binding to DNA targets [5,6,10].

The isolated LIM domains from Lhx3 (Lhx3 LIM1+2) tend to be insoluble and/or aggregate, but soluble stable “tethered complexes” can be engineered in which the LIM interaction domain of Ldb1 (Ldb1 LID), or the Lhx3-binding domain from Isl1 (Isl1 LBD) are fused to Lhx3 LIM1+2 via a flexible glycine-serine linker (Figure 1B) [23,24,25]. These tethered complexes are hereafter referred to as Ldb1-Lhx3 and Lhx3-Isl1 (the order of the names indicates the order of the domains in the complex). Our recently determined structures of Ldb1-Lhx3 and Lhx3-Isl1 show that Ldb1 LBD and Isl1 LBD interact with Lhx3 in an essentially identical manner [13], with the two binding partners forming extended chains and contacting the same sites across both LIM domains of binding protein [17]. The DNA-binding specificity of Lhx3 is context dependent, and varies according to the protein isoform [18,19], or whether Lhx3 is in acting in concert with a protein partner [2,20,21,22]. Disease-causing mutations or post-translational modifications, including phosphorylation, of the LIM domains are likely to affect the biological activity of Lhx3 by modulating protein-protein interactions and modulating binding to DNA targets [5,6,10].
Lhx3. The tethered complexes appeared to have some flexibility in the “hinge” between the LIM domains and the corresponding “spacers” in Ldb1LID and Isl1LBD which lie between the two LIM-binding motifs in each of those domains. The hinge comprises Lhx3F90, which shows some variation in backbone angle in different conformers/molecules, and the spacers comprise Ldb1M310–D318 and Isl1H273–Q278, which assume different overall conformations in the Ldb1 and Isl1 structures (including a short region of disorder in Ldb1 G312–G315 and an extended structure in one chain of Isl1, but a turn in the other chain) [13]. For Lhx3-Isl1 flexibility at the hinge/spacer is supported by small angle X-ray scattering (SAXS) data [26]. The initial part of this study uses SAXS to further characterize the solution structure of Ldb1-Lhx3. These structures provide us with the opportunity to interpret the molecular effects of disease-causing mutations and posttranslational modification of the LIM domains of Lhx3. The mutation of tyrosine 111 to cysteine, Y111C, is an inherited point mutation found in the LIM domains of human Lhx3 that is associated with CPHDS [4]. Although the sequences of the LIM domains from Lhx3 are almost identical in mammals, the numbering of the human and mouse proteins differs slightly; numbering for the mouse protein is used herein, with human Y111 corresponding to Y114 in the mouse protein. The affected residue lies adjacent to one of the zinc ligating residues, H115, prompting suggestions that the introduced cysteine sidechain in Y114C might displace H115 as a zinc ligand [5,6]. Two putative phosphorylation sites, T63 and S71 are located in the first LIM domain of human Lhx3 [10], of these residues only S71 is highly conserved across species. The equivalent residue in the mouse protein, S74, lies adjacent to the binding interface within the Lhx3/Ldb1 and Lhx3/Isl1 complexes.

Here we used SAXS to show that pseudophosphorylation of Lhx3-S74 does not affect binding to Ldb1 or Isl1, but that Lhx3(Y114C) has reduced levels of binding to both key partners. NMR spectroscopy and stability studies demonstrated that Lhx3(Y114C) does not alter the zinc-ligation characteristics of this domain, but does affect the stability and local structure of the second LIM domain of Lhx3.

**Materials and Methods**

**Constructs**

Unless otherwise specified, residue numbers and sequences refer to mouse proteins Lhx3 (UniProtKB/Swiss-Prot ID P50481-1), Ldb1 (UniProtKB/Swiss-Prot ID P70662-3) and Isl1 (UniProtKB/Swiss-Prot ID P61372-1). All constructs were cloned into pGBT9 and pGAD10 (Novagen) for yeast two-hybrid experiments, or pGEX-2T (GE Healthcare) for expression in bacteria with an N-terminal GST (glutathione-S-transferase) tag. All plasmids were sequenced to confirm identity (SUPAMAC, Royal Prince Alfred Hospital, Sydney). The Ldb1-Lhx3 tethered construct contains Ldb1 residues 300–339 (Ldb1 LID), a synthetic linker (GGSGGHMGSGG), and Lhx3 residues 28–153, and were constructed, expressed and purified as described previously [24,25]. Ldb1-Lhx3(Y111C) was generated by PCR using overlap extension mutagenesis. Protein concentrations were determined by absorbance at 280 nm using theoretical extinction coefficients, ε280 nm of 12,210 M⁻¹ cm⁻¹ for wildtype Ldb1-Lhx3 and 10,845 M⁻¹ cm⁻¹ for Ldb1-Lhx3(Y111C).

**Yeast Two-hybrid Assays**

pGBT9 and pGAD10 plasmids were co-transformed into AH109 cells (Clontech), as described previously [27]. All selective media lacked leucine (-L) and tryptophan (-W) to ensure co-transformation of bait and prey plasmids was maintained. For
screening of interactions, all media were further deficient in histidine (-H) but contained or lacked additional reagents for selection of different affinity interactions. Selective media were either supplemented with 40 µg mL⁻¹ X-α-gal (Progen) and 1 mM 3-amino-1,2,4-triazole (3-AT; Sigma) (-L-W-H +3-AT; moderate stringency selection), or were additionally deficient in adenine (-L-W-H-A; high stringency selection).

Circular Dichroism Spectropolarimetry
Far UV-CD experiments were recorded at 20°C on a Jasco J-720 spectropolarimeter equipped with a Neslab RTE-111 temperature controller. Protein samples (30 µM) were prepared in 20 mM Na₂HPO₄, 40 mM NaCl, 1 mM DTT, pH 6.8, placed in a 1-mm path length quartz cell seated in a water-jacketed cell holder. CD spectra were collected over the wavelength range 205–260 nm, a speed of 20 nm/min, 1-nm step resolution, 1-nm bandwidth and a response time of 1 s. Final spectra were the average of five scans, and were buffer baseline corrected. Estimates of secondary structure were determined using CDPro [28].

Chemical Denaturation Experiments
Protein (2.5 µM) in 20 mM Na₂HPO₄, 40 mM NaCl, 1 mM DTT, pH 6.8 and Gdn.HCl as indicated were incubated at 25°C for 2–3 h. Fluorescence emission spectra (320–370 nm) were recorded using a Varian Cary Eclipse fluorescence spectrometer (Palo Alto, CA, USA), with an excitation wavelength of 295 nm. 1 mM 3-amino-1,2,4-triazole (3-AT; Sigma) (-L-W-H +3-AT; moderate stringency selection), or were additionally deficient in adenine (-L-W-H-A; high stringency selection).

Nuclear Magnetic Resonance
NMR samples contained 1 mM 2,2-dimethylsilapentane-5-sulfonic acid (DSS) and 10% D₂O. Spectra were collected at 310 K using a Bruker DRX-600 spectrometer equipped with a 5-mm triple resonance probehead (TXI) and three-axis pulsed field gradients. The spectral widths/carrier frequencies were: 1H - 13 ppm/205 ppm; 15N - 35 ppm/118.3 ppm, or 140 ppm/205 ppm for the detection of histidine sidechains [29].

Data were processed using XWINNMR 3.5 or TOPSPIN 1.2 (Bruker Biospin). Spectral analysis was carried out using SPARKY versions 3.05–3.113 [30]. 1H frequencies of all spectra were directly referenced to DSS, and 15N frequencies were referenced indirectly [31]. Spectral resolution in the directly detected dimension was enhanced by apodisation with either a Lorentzian-Gaussian window function (LB = 0.1, GB = -3) or a squared sine bell function (shifted by π/2.6 radians) in indirectly detected dimensions. Digital resolution was enhanced by zero-filling (once in each dimension) and linear prediction (in the 15N dimension only) before Fourier transformation. Polynomial baseline corrections were applied to processed spectra where appropriate.

Small Angle X-ray Scattering
Small-angle X-ray scattering data were collected from solutions of Ldb1-Lhx3 (4.5–9 mg mL⁻¹) with 20 mM Na₂HPO₄, 40 mM NaCl, 1 mM DTT, pH 6.8 and a matched solvent blank at 283 K (Bruker 200 gm L⁻¹ X-α-gal (Progen) and 1 mM 3-amino-1,2,4-triazole (3-AT; Sigma) (-L-W-H +3-AT; moderate stringency selection), or were additionally deficient in adenine (-L-W-H-A; high stringency selection).

Results
SAXS Data are Consistent with NMR Data for Ldb1-Lhx3
Our previously determined NMR structure of Ldb1-Lhx3 is elongated, with members of the NMR ensemble comprising two LIM modules (LIM1 and LIM2, each with the contacting region of Ldb11-DLIM) but with angles at the “hinge”/“spacer” between the two modules that vary by up to ~30°, and the tether between Ldb1 and Lhx3 and residues at the C-terminus of the construct are unstructured (Figure 2A and B) [13]. Similar tethered complexes (Lhx3-Isi1 and Lhx4-Isi2) can give rise to extreme angles in crystal structures [13,26]. Relatively few long distance restraints exist between the two modules in the NMR structure, and residual dipolar coupling constants could not be determined for this complex [13], creating some uncertainty as to whether the gross structure of the Lhx3/Ldb1 complex was adequately described. Thus we used small-angle X-ray scattering (SAXS) to independently define the global conformation of Ldb1-Lhx3 in solution. Analysis of the molecular weight derived from the forward scattering intensity at zero angle combined with the linearity of the Guinier plots in the low-q regime indicates that Ldb1-Lhx3 exists as a monodisperse sample of monomers (Table 1; Figure 2C) [42].

The experimentally derived P(q) profile (Figure 2D) is characteristic of an elongated molecule [43,44], and is consistent in overall dimensions (~90×20×20; Dmax = 88 Å; Rg = 23.3 Å) with members of the NMR ensemble for Ldb1-Lhx3 (P(q)NMR = 24–25.1 Å; Figure 2A and B, Table S1). The 20 conformers of the NMR ensemble show a range of fits to the SAXS data, with z² values ranging from 0.91 to 1.37 (Text S1, Table S1). The conformer that best fits the data is Model 17 (Figure 2C and D), which still shows a small deviation from the experimental SAXS data (e.g. Figure 2D).

To generate models of Ldb1-Lhx3 based solely on SAXS data, ab initio shape reconstructions of Ldb1-Lhx3, which are independent of high-resolution model bias, were calculated using DAMMIF. Each of the ten individual solutions used to generate
Figure 2. Solution structure of Ldb1-Lhx3. (A) NMR structure of Ldb1-Lhx3 (pdb accession code: 2JTN) with Ldb1 LID in yellow and Lhx3 in blue; zinc ions are depicted as grey spheres. The 20 lowest energy models are aligned over the backbone atoms of both LIM domains. Only the structured regions are shown. The positions of the N and C termini of the interacting domains of Lhx3 and Ldb1, and the position of the two LIM domains from Lhx3 are indicated. (B) Ldb1-Lhx3 with the same 20 conformers from (A) aligned over the backbone atoms of Lhx3 LIM1 and the corresponding region of Ldb1. The unstructured linker and tail from each model is shown in grey. (C) Scattering data for Ldb1-Lhx3 (grey circles) shown as $I(q)$ versus $q$ with the corresponding Guinier plot ($\ln(I(q))$ versus $q^2$) in the inset. The fit is for Model 17 from the Ldb1-Lhx3 NMR ensemble as generated by CRYSOL (black line). The black line in the inset is the fit to the data generated by GNOM. (D) $P(r)$ profiles from experimental scattering data for Ldb1-Lhx3 (grey circles) and calculated scattering profiles from Model 17 of the NMR ensemble (blue line) and the generated BUNCH model (dashed magenta line). (E) Alignment of Model 17 of the Ldb1-Lhx3 NMR ensemble (coloured as in Panel B), BUNCH model (magenta) and the ab initio DAMMIF reconstruction from the scattering data (transparent white surface). (F) Alignment of the NMR ensemble (coloured as in Panel B) with the same DAMMIF consensus model. (G) Most disparate models from the 2JTN NMR structure. Model 1 (blue/yellow) has an angle between the $C\alpha$ atoms of the first zinc-coordinating residue (C34), the hinge residue (F89) and the last zinc coordinating residue (D147) of $163.7^\circ$ (solid black line). Model 9 (cyan/light yellow) has an angle over the same atoms of $135.1^\circ$ (dashed black line). Models are aligned using the backbone atoms of the LIM1 domain. Images of structures were created in Pymol.

doi:10.1371/journal.pone.0040719.g002
the consensus DAMMIF shape of the complex (Figure 2E) fit the experimental data well ($\chi^2$ of 0.7) and have an average normalised spatial discrepancy of 0.63 indicating that the solutions used to generate the consensus model are similar [39]. We also performed rigid-body modelling of Ldb1-Lhx3 using BUNCH [40] in which the positions of the two LIM-modules of the complex were allowed to flex relative to each other during refinement against the SAXS data, as described for related Lhx3/4-Isl1/2 complexes in [26]. The resultant BUNCH model fits the SAXS data with $\chi^2 = 0.75$. Model 17 from the NMR ensemble, the BUNCH model and the DAMMIF shape reconstruction all superimpose well (Figure 2E). The differences between the BUNCH and NMR models are small (Figure 2E), but it should be noted that these proteins have identical levels of secondary structure (6–7% $\alpha$-helical structure, 33–35% $\beta$-structure, and 58–61% coil). The relative stability of the mutant and wildtype tethered complexes was assessed using resistance to chemical denaturation monitored by intrinsic tryptophan fluorescence as described previously [13,26,46]. Both complexes contain a single tryptophan in Lhx3LIM1, and their tryptophan fluorescence emission spectra show a red-shift of wavelength maximum and reduction in intensity typical of unfolding when exposed to 6 M Gdn.HCl (data not shown). Both proteins displayed a monophasic unfolding transition (Figure 4B), but it should be noted that these proteins each contain two LIM domains and a binding peptide meaning

### Table 1. Structural parameters for Ldb1-Lhx3 from small-angle X-ray scattering data.

| Parameter       | Value       |
|-----------------|-------------|
| $D_{\text{max}}$ (Å) | 88          |
| [protein] (mg mL$^{-1}$) | 4.50 ± 0.12$^a$ |
| (f0) (cm$^{-1}$) | 0.0890 (±0.0004)$^a$ |
| $\Delta\rho$ (x10$^{15}$ cm$^{-2}$) | 2.29        |
| Rg (Å)          | 25.3 (±0.2)$^a$ |
| Experimental molecular mass (MM$_{exp}$) (kDa) | 22.0 ± 0.6 |
| Predicted molecular mass (MM$_p$) (kDa) | 20.5 |
| MM$_{exp}$/MM$_p$ | 1.07        |

We use the convention for reporting SAXS data as outlined in ref [42].

$^a$error is 1 S.D. 

$^b$MM$_{exp}$ = f0/Nu/[protein]A$_0\rho^2$.

*Model 17 and $R_g$ were derived from P(r) using GNOM.

doi:10.1371/journal.pone.0040719.t001

Analysis of Interactions between Lhx3 Mutants

We used yeast two-hybrid analysis to test if the Lhx3(Y114C) mutant and two pseudo phosphorylation mutants [45], Lhx3(S74D) and Lhx3(S74E), affected the interaction of that protein with Ldb1LID or Isl1LBD (Figure 3). Yeast two-hybrid data for Lhx3 (constructs comprised both LIM domains), the Lhx3 phosphomimic mutants Lhx3(S71D) and Lhx3(S71E) and the CPHDS mutation Lhx3(Y114C) against Ldb1LID and Isl1LBD, and interaction data for both bait/prey orientations are shown; DBD and AD designate the plasmids pGBT9 and pGAD10, respectively. Serial dilutions of culture ($A_{600nm}$ of 0.2, 0.02, and 0.002) were spotted onto each column of the plate with the highest concentration at the top. Two different selection conditions (moderate stringency = SD-L-W-H-A; high stringency = SD-L-W-H-A) are represented. Data are representative of 2-3 separate experiments. In all cases transformation control plates (SD-L-W-H-A) showed strong growth of yeast indicating successful transformations, and negative controls of each constructs versus the corresponding empty plasmid showed no yeast growth under the conditions tested.

doi:10.1371/journal.pone.0040719.g003

Figure 3. Effect of mutations in Lhx3 on interactions with Ldb1 and Isl1. Yeast two-hybrid data for Lhx3 (constructs comprised both LIM domains), the Lhx3 phosphomimic mutants Lhx3(S71D) and Lhx3(S71E) and the CPHDS mutation Lhx3(Y114C) against Ldb1LID and Isl1LBD, and interaction data for both bait/prey orientations are shown; DBD and AD designate the plasmids pGBT9 and pGAD10, respectively. Serial dilutions of culture ($A_{600nm}$ of 0.2, 0.02, and 0.002) were spotted onto each column of the plate with the highest concentration at the top. Two different selection conditions (moderate stringency = SD-L-W-H-A; high stringency = SD-L-W-H-A) are represented. Data are representative of 2-3 separate experiments. In all cases transformation control plates (SD-L-W-H-A) showed strong growth of yeast indicating successful transformations, and negative controls of each constructs versus the corresponding empty plasmid showed no yeast growth under the conditions tested.

**Effect of mutation on interaction**

- $+$ no change
- +/- moderate
- - strong

| Ldb1 | Isl1 |
|------|------|
| Lhx3-S74D | +  | + |
| Lhx3-S74E | +  | +/–|
| Lhx3-Y114C | +/-| – |
that the folding was unlikely to be two-state or fully reversible and free energies of folding could not be determined. The mutant tethered complex (midpoint of denaturation, $D_{50\%} = 2.0 \text{ M Gdn.HCl}$) was apparently destabilised relative to the wildtype complex ($D_{50\%} = 2.5 \text{ M Gdn.HCl}$). The stability of these complexes stems from both the intrinsic stability of the LIM domains and the affinity of the LIM domains for Ldb1-LID, and so the decrease in the overall apparent stability of the Ldb1LIM-Lhx3(Y114C) mutant likely reflects a weaker interaction and/or decrease in stability of the Lhx3LIM2.

The structure of the mutant protein was further probed by NMR spectroscopy. The peaks in the $^{15}$N-HSQC spectrum of uniformly labelled $^{15}$N-Ldb1-Lhx3(Y114C) were sharp and well dispersed indicating that the protein is folded; however, many of the peaks have shifted compared to the spectrum for the wildtype protein (Figure 4C); we have inferred assignments for the mutant protein from resonance proximity in two-dimensional spectra (Figure S2 and Table S3). Although caution should be taken in the interpretation of these spectra because we have not unambiguously assigned the spectra, the majority of the peaks that have moved in the spectrum of the mutant protein correspond to residues in the Lhx3LIM2 domain whereas the N-terminal half of Ldb1LID and Lhx3LIM1 appear to be largely unaffected by the mutation (Figure 4D, Figure S2 and Table S3).

We analysed the protonation pattern of the H115 sidechain using $^{15}$N-HSQC experiments [29] to investigate the possibility that the Y114C mutation might alter the zinc coordination. The two histidine residues that ligate zinc ions in Ldb1-Lhx3, Lhx3(H55) and Lhx3(H115), give rise to patterns that are typical of protonation at $N^\text{e2}$, indicating that zinc ligation occurs through $N^\text{d2}$ in both cases (Figure 5). Any change in the zinc coordination state of those histidine sidechains would result in changes to both the pattern and intensity of the peaks in the spectrum of the mutant protein (Figure 5A). Although there are some changes in the $^{15}$N chemical shift resonances from the $N^\text{d2}$ nitrogen of H115 in the spectrum from Ldb1-Lhx3(Y114C), the lack of movement of the resonances corresponding to the $N^\text{e2}$, $H^\text{e1}$ and $H^\text{d2}$ nuclei suggests that the ligation state of H115 is the same in both the wildtype and mutant proteins (Figure 5B). That is, C114 in the mutant protein does not replace H115 as a zinc ligand. The movement of the $N^\text{d1}$ nitrogen of H115 can be attributed to its proximity to the site of mutation; a bulky aromatic tyrosine sidechain has been substituted by cysteine, likely causing a change in the local fold and/or electronic structure of the protein.

**Figure 4. Effect of mutations on Lhx3.** (A) Overlay of CD spectra from Ldb1-Lhx3 (solid black squares) and Ldb1-Lhx3(Y114C) (open magenta circles). Spectra were collected with 30 μM samples in 20 mM Na$_2$HPO$_4$, 40 mM NaCl, 1 mM DTT, pH 6.8 at 310K and were buffer baseline corrected. (B) Chemical denaturation of Ldb1-Lhx3 (solid black squares) and Ldb1-Lhx3(Y114C) (open magenta circles). Fraction folded was estimated using tryptophan fluorescence intensities (excitation wavelength 295 nm and emission wavelength 334 nm). Lines show the fits to a sigmoidal function. Proteins were at concentrations of 2.5 μM. (C) $^{15}$N-HSQC spectra of Ldb1-Lhx3 (black, ~800 μM) and Ldb1-Lhx3(Y114C) (magenta, ~200 μM) in 20 mM Na$_2$HPO$_4$, 40 mM NaCl, 1 mM DTT, pH 6.8 at 310K. (D) Analysis of chemical shift differences from panel C based on assignments for the wildtype protein [25] and inferred assignments for the mutant protein (Table S3). Peaks were identical for the C-terminal half of Ldb1LID and Lhx3LIM1 which make direct contacts in the structure of the complex, but were significantly different for the N-terminal half of Ldb1LID and Lhx3LIM2.
Discussion

The SAXS analysis of Ldb1-Lhx3 confirms that the previously published NMR structure of this tethered complex is a reasonable representation of its solution structure - an ensemble of conformers that varies in angle by up to $30^\circ$ between the two LIM modules. Given the extended nature of this complex and the paucity of contacts between the two LIM modules we expect that the complex undergoes a limited amount of flexion at this point and/or twists as suggested by the BUNCH model (Figure 2). The SAXS data for Ldb1-Lhx3 are consistent with a small amount of flexion, but not with high levels of flexibility. Previously reported heteronuclear $^{15}\text{N}-^1\text{H}$ NOE values are consistent with no or limited flexion between the LIM domains, although increased motion in residues Ldb1(E313–G315) corresponds with a short region of disorder in Ldb1LID spacer [13]. Flexibility at the hinge/
spacer was suggested for a related LIM-only protein 2 (Lmo2)-Ldb1LID complex [47], and is consistent with SAXS data for Lhx3/4-Isl1/2 complexes [26]. A phenylalanine residue followed by a glycine residue is highly conserved at this hinge in LIM-homeodomain and LIM-only proteins (Figure S3), suggesting that some flexion at the hinge/spacer is a common phenomenon in this type of complex. Flexibility was suggested to play a role in the binding of Lmo2/Ldb1 to Tal1/E47 [47], and differences in the inter-LIM domain angle could influence the recruitment of additional binding partners to and thereby generate differences in activity between alternate Lhx3-containing complexes.

We recently showed that the complex formed between LIM homeobox protein 4 (Lhx4) and Islet 2 (Isl2) and that formed between Lhx3 and Isl1 have a more compact average structure (Dmax ~75 Å) than a similar complex formed by Lhx3 and Isl1 (Dmax ~90 Å), suggesting that the binding partner can influence the average angle formed between the two LIM domains [26]. Despite a higher sequence identity between Isl1LBD and Isl2LBD than Isl1LBD and Ldb1LID, SAXS data for Ldb1-Lhx3 [which show, for example, a maximum dimension Dmax of 88 Å] indicate that the gross structure of this complex resembles the more elongated Lhx3/4-Isl1 complexes. The spacing and structure of spacer may play an important role in determining the inter-LIM angle; the spacers in Isl1LBD and Isl2LBD have very low sequence identity compared to the LIM-binding motifs in those domains.

Although S74 lies proximal to the binding faces in both complexes (Figure 6), neither of the phosphomimetic mutants of Lhx3 had a major effect on binding to Ldb1 and Isl1 using yeast two-hybrid assays. Only S74E showed some apparent reduction in binding suggesting that the larger size (rather than the charge) of the glutamate sidechain was responsible for this effect. It is plausible that in Lhx3/S74E local rearrangements in the LIM1 domain may be required to accommodate the larger glutamate sidechain resulting in a minor reduction of binding. These results are consistent with data from Parker et al. showing that Lhx3/S74A mutants did not affect binding of Lhx3 to Ldb1, PIT-1 or MRG in GST-pulldown experiments [10]. If Lhx3 is phosphorylated at this site as part of normal activity it may be to modulate binding to as yet unidentified or uncharacterised protein partners.

Our yeast two-hybrid data using the CPHDS-associated mutation Lhx3(Y114C) confirm the reduced binding observed between Lhx3 mutants and Ldb1 in GST-pulldown experiments [6]. That the mutation also severely affects interactions with Isl1 is consistent with our structural and mutagenic scanning data showing that Isl1 and Ldb1 interact with the same site on Lhx3, with the majority of critical contacts being made at the LIM2 domain [13]. Y114 is highly conserved in the LIM domains of LIM-homeodomain and LIM-only proteins (Figure S3) suggesting that Y114 is crucial for the structure and/or function of these proteins. Y114 does not contact either Ldb1LID or Isl1LBD (Figure 6A), nor does the main and LIM-only proteins (Figure S3) suggesting that Y114 is highly conserved in the LIM domains of LIM-homeodomain and LIM-only proteins (Figure S3), suggesting that Y114 is highly conserved in the LIM domains of LIM-homeodomain and LIM-only proteins (Figure S3). Y114 is highly conserved in the LIM domains of LIM-homeodomain and LIM-only proteins (Figure S3).

Our NMR structure of Ldb1-Lhx3 (pdb accession code: 2JTN) with 8 (magenta linker), 9 (orange linker) and 17 (black linker) from the top section of the table, each of the NMR models was fitted as indicated with a dash. Residues equivalent murine Lhx3 Y114 are indicated models from the Ldb1-Lhx3 NMR ensemble as indicated. Note that the scale on the Y-axis is arbitrary and the data/fits have been plotted separately for clarity.

In conclusion, our analysis of Lhx3 complexes provides a structural framework with which to understand how Lhx3 is regulated in the cell, and how the biological functions of the protein may be affected by disease causing mutations.

Supporting Information

Figure S1 Analysis of SAXS data. (A) Models 1 (cyan linker), 8 (magenta linker), 9 (orange linker) and 17 (black linker) from the NMR structure of Ldb1-Lhx3 (pdb accession code: 2JTN) with Ldb1LID in yellow and Lhx3 in blue; zinc ions are depicted as grey spheres. The models were aligned over the backbone atoms of Lhx3LIM2 and the corresponding region of Ldb1. (B) Scattering data for Ldb1-Lhx3 (grey circles) shown as I(q) against q plots. Curves are the fits to the data generated by CRYSOL of sample NMR models from the Ldb1-Lhx3 NMR ensemble as indicated. Note that the scale on the Y-axis is arbitrary and the data/fits have been plotted separately for clarity.

Figure S2 Overlay of 15N-HSQC spectra for wildtype and mutant Ldb1-Lhx3 constructs. 15N-HSQC spectra of Ldb1-Lhx3 [black; 800 μM protein] and Ldb1-Lhx3(Y114C) [red; ~200 μM protein]. Spectra were collected at 310 K in 20 mM Na2HPO4, 40 mM NaCl, 1 mM DTT, pH 6.8. Assignments for wildtype Ldb1-Lhx3 [25] are shown in black; for clarity, not all assignments are shown. Note that construct residues 1–45 refer to Ldb1295–339; construct residues 46–56 correspond to the synthetic linker, and construct residues 57–182 refer to Lhx328–153 (See Table S3 for full list).

Table S1 Fits to the SAXS data for the individual models from the NMR models. The fits are reported as χ2 values. For the top section of the table, each of the NMR models was fitted as the intact molecule, or with the unstructured tether and tail regions removed.

Table S2 Fits to the SAXS data for “swapped” tethers. The chimera was generated by swapping the linkers between the indicated models. Value reported are χ2 values of the fit.

Table S3 Inferred assignments of resonances in the 15N-HSQC of Ldb1-Lhx3(Y114C). Values are given for the wildtype assignments [25], the inferred assignments for the Y114C
mutant and the weighted average chemical shift differences between those assignments. Note that not all peaks could be assigned by this approach. The shaded parts of the table refer to assignments in the N-terminal half of Ldb1 (LIM and Lhx3) (DOCX).

Text S1 Fits of NMR structures and SAXS data. (DOCX)

References
1. Sheng HZ, Zhadanov AB, Mosinger B Jr, Fuji I, Bertuzzi S, et al. (1996) Specification of pituitary cell lineages by the LIM homeobox gene Lhx1. Science 272: 1004–1007.
2. Thaler JP, Lee SK, Jurata LW, Gill GN, Pfaff SL (2002) LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. Cell 110: 237–249.
3. Sharma K, Sheng HZ, Lettieri K, Li H, Karavanov A, et al. (1998) LIM homeodomain factors Lhx3 and Lhx1 assign subtype identities for motor neurons. Cell 93: 817–828.
4. Nethcine I, Sobrier ML, Krude H, Schnabl D, Maghnie M, et al. (2000) Mutations in LHX3 result in a new syndrome revealed by combined pituitary hormone deficiency. Nat Genet 25: 182–186.
5. Sloep KV, Parker GE, Hanna KR, Wright HA, Rhodes SJ (2001) LHX3 transcription factor mutations associated with combined pituitary hormone deficiency impair the activation of pituitary target genes. Gene 263: 61–69.
6. Howard PW, Maurer RA (2001) A point mutation in the LIM domain of Lhx3 reduces activation of the glycoprotein hormone alpha-subunit promoter. J Biol Chem 276: 19020–19026.
7. Pfaffle RW, Savage JJ, Hunter CS, Palme C, Altmann M, et al. (2007) Four novel mutations of the LHX3 gene cause combined pituitary hormone deficiencies with or without limited neck rotation. J Clin Endocrinol Metab 92: 1909–1919.
8. Savage JJ, Hunter CS, Clark-Sturm SL, Jacob TM, Pfaffle RW, et al. (2007) Mutations in the LHX3 gene cause dysregulation of pituitary and neural target genes that reflect patient phenotypes. Gene 400: 44–51.
9. Rajab A, Kelberman D, de Castro SCP, Biebermann H, Shaikh H, et al. (2008) Novel mutations in LHX3 are associated with hypopituitarism and sensorineural hearing loss. Hum Mol Genet 17: 2150–2159.
10. Parker GE, West BE, Witzmann FA, Rhodes SJ (2005) Serine/threonine-proximal phosphorylation of the LHX3-LIM-homeodomain transcription factor. J Cell Biochem 94: 67–80.
11. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
12. Jurata LW, Pfaff SL, Gill GN (1998) The nuclear LIM domain interactor NLI mediates homo- and heterodimerization of LIM domain transcription factors. J Biol Chem 273: 3152–3157.
13. Bhati M, Lee C, Nancarrow AL, Lee M, Craig VJ, et al. (2008) Implementing the LIM code: the structural basis for cell-type-specific assembly of LIM-domain complexes. EMBO J 27: 2019–2029.
14. Agulnick AD, Taira M, Breen JJ, Tanaka T, Davold IB, et al. (1996) Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins. Nature 384: 270–273.
15. Meier BC, Price JR, Parker GE, Bridwell JL, Rhodes SJ (1999) Characterization of the porcine Lhx3/LIM-3/P-Lim LIM homeodomain transcription factor. Nature 384: 270–272.
16. Bhati M, Lettieri K, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
17. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
18. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
19. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
20. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
21. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
22. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
23. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.
24. Matthews JM, Bhati M, Lehtomaki E, Mannfield RE, Czubeka L, et al. (2009) It takes two to Tango: the structure and function of LIM, RING, PHD and MYND domains. Current Pharmaceutical Design 15: 3681–3696.