Supporting Information

Wei et al. 10.1073/pnas.1016567108

SI Materials and Methods

Protein Expression and Purification. The coding sequence of MyoX_MF (NCBI access number: NP_036466, residues 1503–2047) was PCR amplified from a human brain cDNA library and cloned into a pET32a vector. The coding sequences of DCC_P3 (NCBI access number: NP_036973, residues 1409–1445) and its N-terminal extension versions (residues 1321–1445 for P2–P3 and 1370–1445 for the extended P3) was PCR amplified from rat *DCC* gene. Its various mutants were generated by PCR based site-directed mutagenesis. The coding sequences for neogenin_P3 (residues 1397–1477) was PCR amplified from mouse *neogenin* gene. For MyoX_MF/DCC_P3 fusion constructs, DCC_P3 (residues 1409–1445) was fused to MyoX_MF at its N terminus or C terminus by PCR. The C-terminal fusion protein has two linker residues (Ser and His) between MyoX_MF and DCC_P3 introduced by the cloning process. DCC fragments and their various mutants were expressed as either His_6-tagged or GST-tagged proteins and purified using Ni^{2+}-nitrilotriacetic acid agarose affinity chromatography or with GSH-Sepharose column, followed by another step of gel filtration chromatography. Fusion proteins were expressed as His_6-tagged proteins and purified using Ni^{2+}-nitrilotriacetic acid agarose affinity chromatography, followed by one more step of gel filtration chromatography.

Crystallography. Crystals of the MyoX_MF/DCC_P3 fusion protein were obtained by hanging drop vapor diffusion method at 16 °C within 2 days. To set up a hanging drop, 1 μl of concentrated protein solution (10 mg/ml) was mixed with 1 μl of crystallization solution with approximately 8% PEG8000 and 10% glycerol in 0.1 M HEPES buffer (pH 7.5). To prepare heavy-atom derivatives, crystals were soaked in the crystallization solution containing 1 mM $KAu(CN)_2$ for two days. Before diffraction experiments, crystals were soaked in crystallization solution containing additional 10% glycerol for cryoprotection. Several diffraction datasets of native crystal and its Au-derivative were collected at 100 K on a Rigaku RAXIS IV++ imaging-plate system with a MicroMax-007 copper rotating-anode generator. The diffraction data were processed and scaled using the MOSFLM and SCALA in the CCP4 suite (1).

Four Au sites were found by SHELXD (2). The site refinements and phase improvements were carried out by autoSHARP (3). After manual backbone building, the phase was further improved by RESOLVE (4) and then used as input for ARP/wARP model building (5). The initial model was refined in Refmac5 (6) against the 2.5-Å native dataset. COOT (7) was used for model rebuilding and adjustments. In the final stage, an additional TLS refinement was performed in Refmac5 (8). The final model was also used for the refinement of a 2.7-Å dataset, which was collected from a crystal in the same crystallization condition as the 2.5-Å one but containing 50 mM NaI.

Structure and Sequence Analysis. The sequence alignment of the MyTH4 domain combined both the primary sequences of the domains as well as the structure of the MyoX MyTH4 structure determined in this work. The consensus pattern of the alignment sequences was characterized using WebLogo (9). All structure figures were prepared by PyMOL (http://www.pymol.org/).

GST Pulldown Assay. The plasmid coding the GFP-tagged full-length human MyoX expression was a gift from Wencheng Xiong. The point or deletion mutations of MyoX: MyoX_MF (residues 1503–2047), MyoX_MyTH4 (residues 1503–1703), MyoX_FERM (residues 1696–2047), SH mutant (residues 1503–2047, S1718 H1719/AA), SHDEK mutant (residues 1503–2047, S1718 H1719 D1763 E1769 K1770/AAAAA) were cloned into pEGFP-C-3 vector for GFP-tagged protein expression. Direct interactions between DCC_P3 and various MyoX MyTH4-FERM mutants were assayed in phosphate-buffered saline (pH 7.4). GST-DCC_P3 fragment (approximately 0.6 nmol each) was first incubated for 30 minutes with GSH-Sepharose beads, and the GST-DCC_P3 loaded beads washed twice each with 0.5 ml assay buffer. GFP-tagged MyoX_MF, MyoX_MyTH4, MyoX_FERM, MyoX_MF(SH/AA), and MyoX_MF(SHDEK/AAAAA) were transiently transfected in HEK293T cells with lipofetamine PLUS kit (Invitrogen) and incubated for 18 hr at 37 °C. Cells were lysed in 1% Triton X-100 lysis buffer (150 mM NaCl, 10% glycerol, 50 mM HEPES pH7.6, 1.5 mM $MgCl_2$, 0.1 M NaF, 1 mM EGTA) on ice for 30 min and centrifugated at 13,000 rpm for 30 min. Lysate was each pelleted by the GST-DCC_P3 coated GSH-Sepharose beads. The pellets were washed three times each with 0.5 ml of the assay buffer, subsequently boiled with 15 μl of 2× SDS/PAGE sample buffer and separated by SDS/PAGE. The GFP-tagged proteins were visualized by immuno-detection using antiGFP antibody.

Isothermal Titration Calorimetry Assay. ITC measurements were carried out on a VP-ITC calorimeter (Microcal) at 25 °C. All protein samples were in 50 mM Tris buffer, pH 7.5 containing 100 mM NaCl. The titration processes were performed by injecting 10 μl aliquots of the Trx-tagged DCC_P3 or neogenin_P3 fragments into MyoX_MF proteins at time intervals of 3 minutes to ensure that the titration peak returned to the baseline. The titration data were analyzed using the program Origin7.0 from Microcal and fitted using the one-site binding model.

Sedimentation Equilibrium. The MyoX_MF/DCC_P3 fusion protein with different concentrations was analyzed on a Beckman XL-I analytical ultracentrifuge under 11,000 rpm at 25 °C. The partial specific volume of protein sample and the solvent density were calculated using the program SEDNTERP (http://www.rasmb.bbri.org/). The final sedimentation equilibrium data were analyzed and fitted to a single-species model to get the molecular mass using the XL-A/XL-I data analysis software provided by the manufacture.

1. Collaborative Computational Project N (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
2. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D* 58:1772–1779.
3. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–30.
4. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D* 56:965–972.
5. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6:458–463.
6. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
7. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
8. Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* 374:300–321.
9. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188–1190.

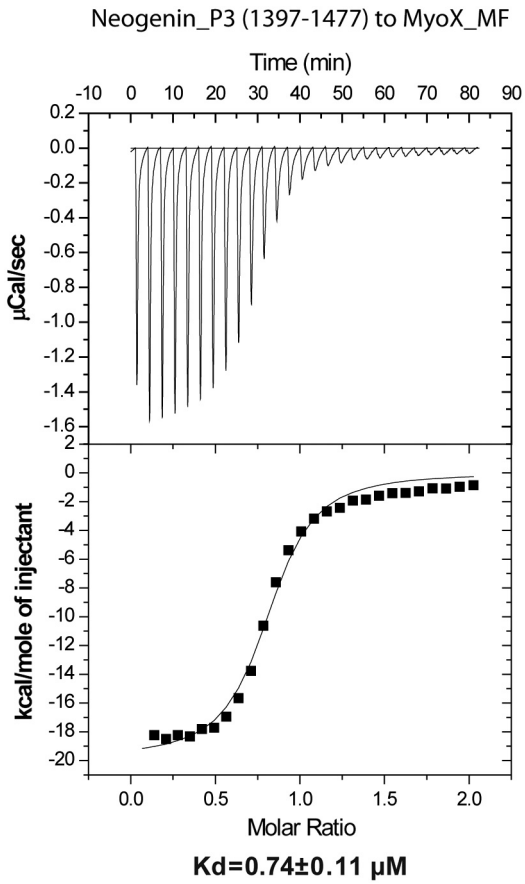

Fig. S1. The ITC analysis of the interactions between MyoX_MF and Neogenin_P3. The details of the experiments are identical to that in Fig. 1*C*.

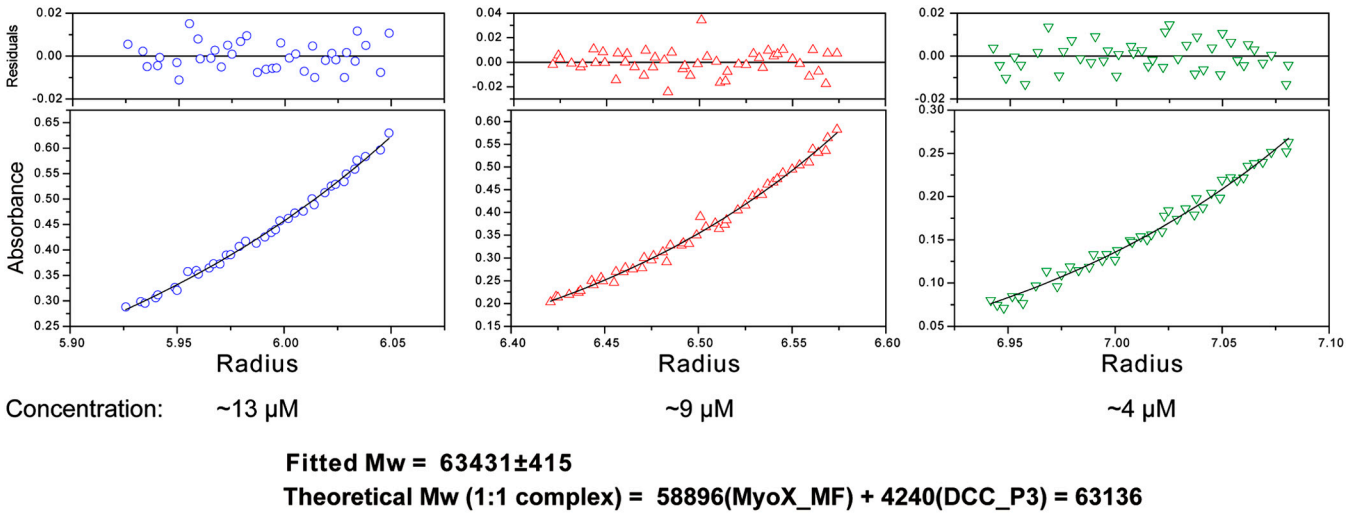

Fig. S2. Sedimentation equilibrium analysis of MyoX_MF/DCC_P3. The figure shows the sedimentation profiles of the MyoX_MF/DCC_P3 fusion protein at three different concentrations. The rotor speed for the sedimentation experiment was 11,000 rpm. The theoretical molecular weight was calculated from its amino acid sequence.

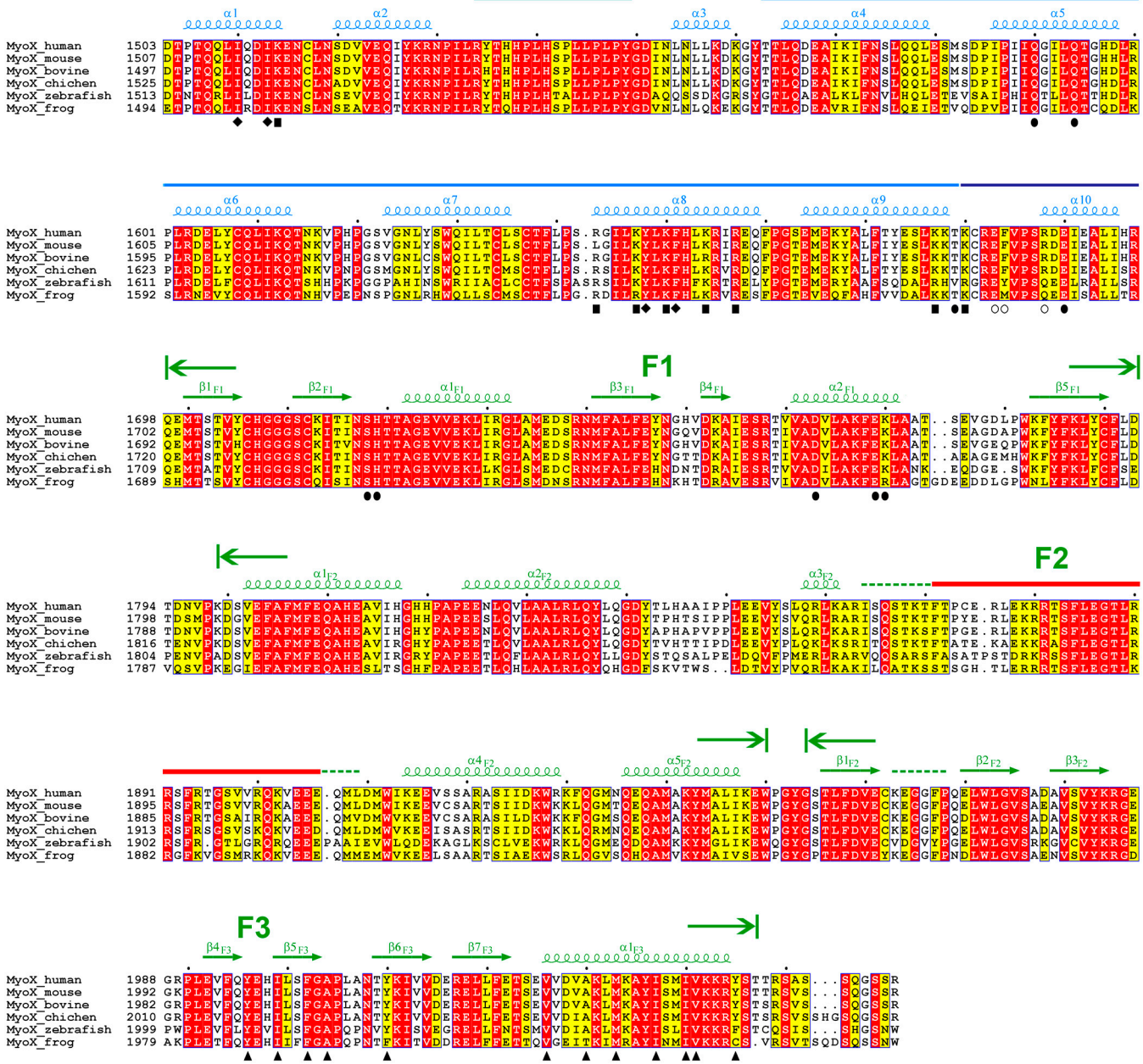

Fig. S3. Sequence alignment of MyoX_MF from different species. Residues that are absolutely and highly conserved are shown in red and yellow boxes, respectively. The secondary structure elements are labeled above the alignment. The disordered regions in the structure are indicated by dashed lines. The 36-residue deletion in the α3/α4-loop of the F2 lobe is indicated by a red line. The boundaries of the MyTH4 domain (residues 1503–1697) and the three lobs of FERM are also indicated. The different segments in the MyTH4 domain are indicated by different color lines above the secondary structure elements (also see Fig. S4). Residues in the F3 lobe that are involved in the binding to DCC_P3 are indicated by triangles. Residues forming interdomain sidechain–sidechain and sidechain–mainchain hydrogen bonds in the MyTH4/F1 interface (see Fig. 6) are indicated by solid and open circles, respectively. Residues involved in forming the positively charged surface and the hydrophobic pocket on the MyTH4 (Fig. 3C and Fig. S7) are indicated by squares and diamonds, respectively.

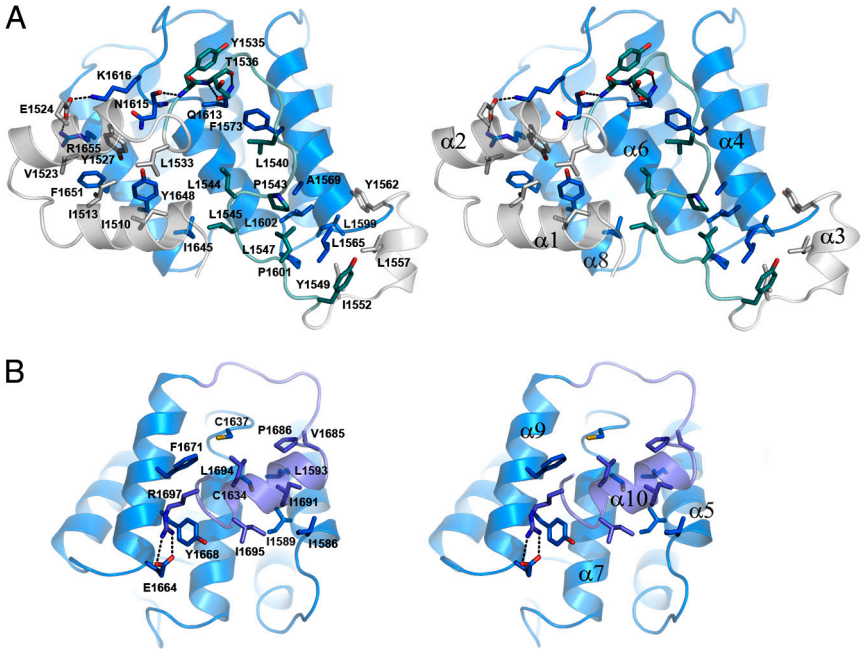

Fig. S4. Interactions between different segments within MyoX MyTH4. The MyTH4 domain is displayed from top (*A*) and bottom (*B*) in stereo view. The first two helices, the α2/α3 loop, the six-helix bundle, and the last helix (α10) are colored in white, green, blue, and purple, respectively. Hydrogen bonds and salt bridges are indicated by dashed lines.

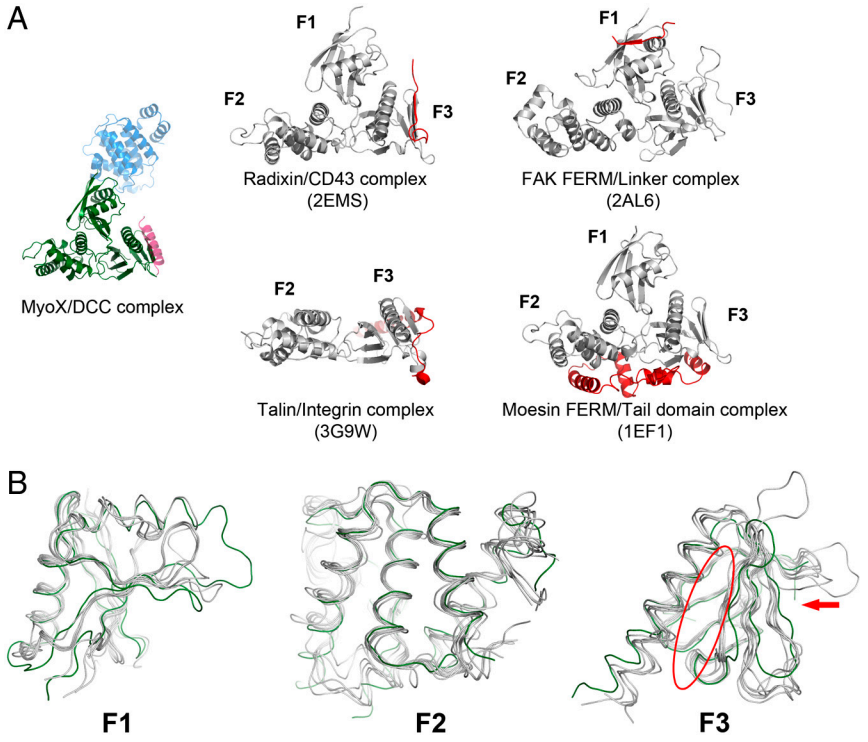

Fig. S5. Comparison of the structure of MyoX FERM with those of other FERM domains. (*A*) Ribbon representations of the FERM domains from Moesin, Talin, Radixin, and FAK each with the similar orientation to that of MyoX FERM. Corresponding PDB ID code for each FERM domain is indicated in the parenthesis. The available structure of the Talin FERM domain only contains F2 and F3 lobes. The bound peptides of the FERM domains are highlighted by red colored ribbons. (*B*) Superposition of the individual lobes of the FERM domains. The figure reveals that the individual lobes of various FERM domains have similar conformations. The αβ- and ββ-grooves in the F3 lobe are indicated by a red oval and an arrow, respectively.

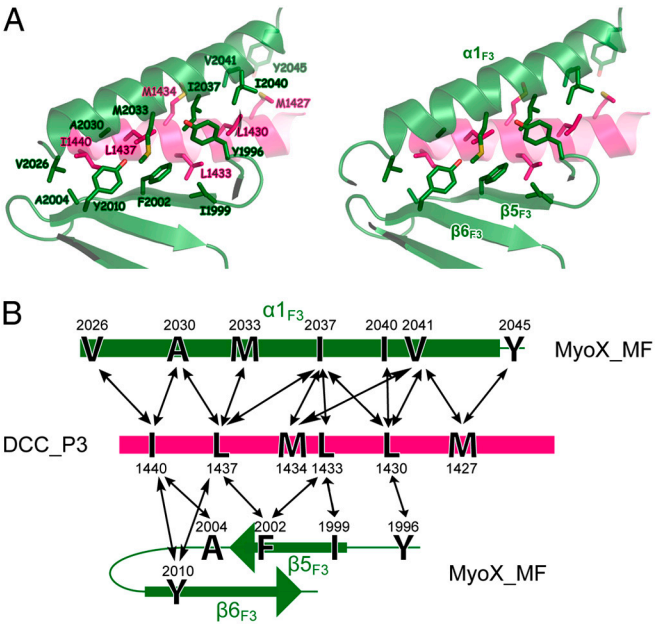

Fig. S6. The hydrophobic interaction between MyoX-MF and DCC is shown as stereo structural view (*A*) and cartoon representation (*B*). The hydrophobic interactions between the side chains of the αβ-groove residues of MyoX FERM F3 lobe (green) and that of the DCC_P3 residues (red) were indicated by two-way arrows in *B*.

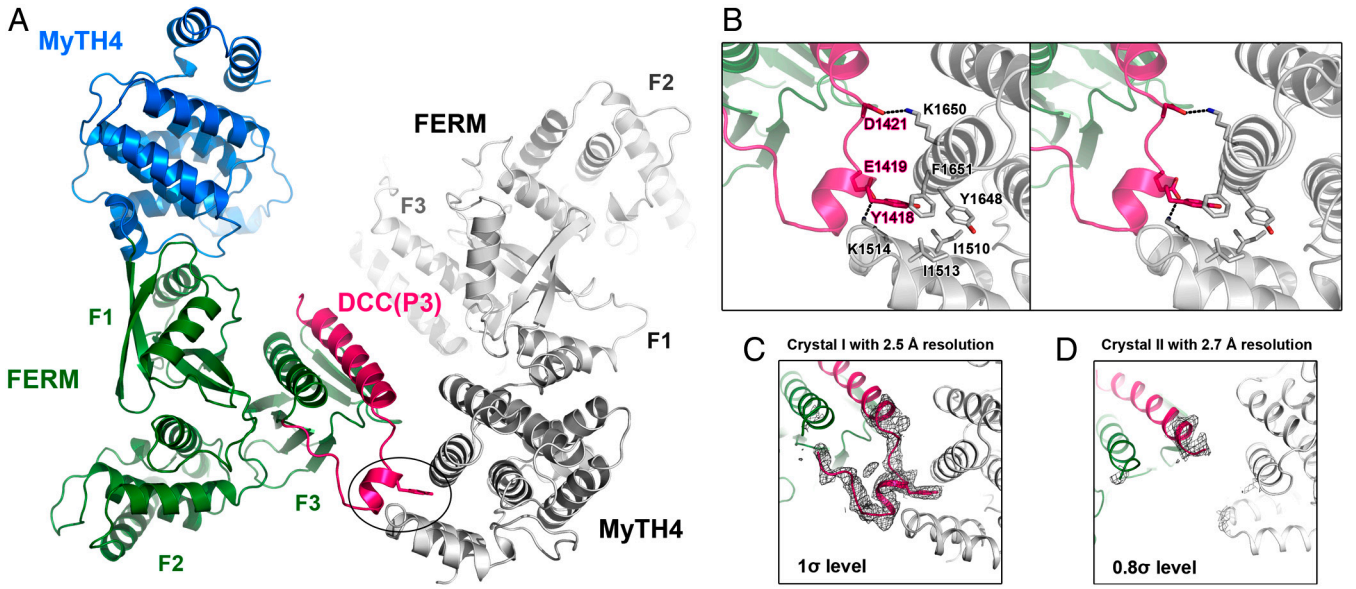

Fig. S7. The intermolecular MyTH4/DCC interaction induced by crystal packing. (*A*) A few residues in the N-terminal end of DCC_P3 forms a small α-helix, which is stabilized by interacting with a symmetric related neighboring MyTH4 domain (gray color) in crystal. Tyr1416 in the small helix inserts its side chain into a pocket on the MyTH4 domain. (*B*) Stereo view of enlarged molecular details of the MyTH4/DCC interaction. Two salt bridges sandwiching $Tyr1416_{DCC}$ are indicated by dashed lines. (*C* and *D*) The difference in the crystal packing between MyTH4 and DCC_P3 in two different crystals with the same space group. Electron densities around the N-terminal short α-helix in the 2.5-Å structure (contoured at 1σ, *C*) and the corresponding region in the 2.7-Å structure (contoured at 0.8σ, *D*) are shown by meshes. The figure demonstrates that the formation of the short N-terminal α-helix in the DCC_P3 is likely to be an artifact induced by crystal packing.

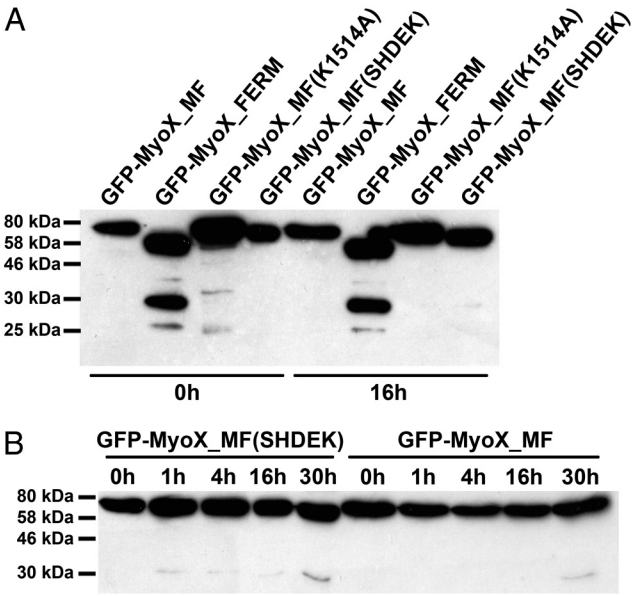

Fig. S8. Protein stabilities of MyoX_MF and its mutants expressed in HEK 293T cells. Human kidney 293T cells were cultured in DMEM supplemented with 10% FBS. N-terminal GFP-tagged proteins were transiently transfected by lipofetamine reagents (Invitrogen). Cells were lysed in 1% Triton X-100 lysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.1 mM NaF, 1 mM EDTA, 10% glycerol) on ice without any additional protease inhibitors for 30 min and centrifuged at 13,000 rpm for 30 min. The supernatants were incubated at 4 °C for various lengths of time to evaluate the stabilities of various forms of soluble MyoX_MF proteins expressed in HEK293T cells. In *A*, samples were collected right after centrifugation (0 hour) and after 16-hour incubation and then were separated by SDS/PAGE. The expressed proteins were immuno-detected using antiGFP antibody. *B* shows a comparison of the sample stabilities of the wild-type MyoX_MF and the MyoX-MF(SHDEK) mutant after incubating each of the cell lysate mixture for the indicated time periods.

Table S1. Statistics of data collection and model refinement

Data collection		
Data sets	Native	Au-derivative
Space group	$P2_1$	
Unit cell parameters (Å)	$a = 85.2$, $b = 49.5$, $c = 92.8$ $\beta = 112.6$	$a = 87.1$, $b = 49.7$, $c = 91.5$ $\beta = 115.9$
Resolution range (Å)	50–2.5 (2.64–2.5)	50–3.0 (3.16–3.0)
No. of unique reflections	24981 (3589)	14470 (2109)
Redundancy	3.7 (3.7)	10.6 (10.3)
Redundancy	3.7 (3.7)	10.6 (10.3)
I/σ	16.8 (2.7)	16.9 (7.5)
Completeness (%)	99.5 (99.1)	100.0 (100.0)
R_{merge} (%) *	7.3 (46.8)	10.8 (44.7)
R_{meas} (%) †	8.6 (54.7)	11.3 (47.0)
Structure refinement		
Resolution (Å)	30–2.5 (2.564–2.5)	
Rcryst / R-free (%) ‡	20.3 (26.4)/25.9 (30.8)	
rmsd bonds (Å) / angles (°)	0.008 / 1.15	
Average B factor	50.6	
No. of atoms		
Protein atoms	4220	
Water molecules	101	
Other molecules	6	
No. of reflections		
Working set	23737	
Test set	1243	
Ramachandran plot regions §		
Favored (%)	97.1	
Allowed (%)	99.6	
Outliner (%)	0.4	

Numbers in parentheses represent the value for the highest resolution shell.

*$R_{merge} = \sum |I_i\text{-}I_m| / \sum I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of all symmetry related reflections.

†R_{meas} is the multiplicity-weighted R factor (10).

‡$R_{cryst} = \sum ||F_{obs}|\text{-}|F_{calc}|| / \sum |F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors. $R_{free} = \sum_T ||F_{obs}|\text{-}|F_{calc}|| / \sum_T |F_{obs}|$, where T is a test dataset of about 5% of the total reflections randomly chosen and set aside prior to refinement.

§Defined by MolProbity (11).