Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification and crystallization of the N-terminal domain from the human doublecortin-like kinase

Myung Hee Kim,[a] Urszula Derewenda,[a] Yancho Devedjiev,[a] Zbigniew Dauter[b] and Zygmunt S. Derewenda[a]*

[a]Department of Molecular Physiology and Biological Physics and the Cancer Center, University of Virginia, PO Box 800736, Charlottesville, Virginia 22908-0736, USA, and [b]Synchrotron Radiation Research Section, Macromolecular Crystallography Laboratory, NCI, Brookhaven National Laboratory, Upton, New York 11973, USA

Correspondence e-mail: zsd4n@virginia.edu

Received 30 August 2002
Accepted 2 January 2003

The unique doublecortin-like tandem of two homologous domains is found in certain microtubule-associated proteins such as doublecortin (DCX) and doublecortin-like kinase (DCLK). It is responsible for interactions with tubulin/microtubules and regulates microtubule dynamics. Here, the expression and purification of the tandem from human DCLK (residues 49–280) and of the isolated domains (residues 49–154 and 176–280) and the successful crystallization of the N-terminal domain (N-DCLK) are reported. High-quality wild-type crystals were obtained and a complete native data set was collected to 1.5 Å resolution. The crystals belong to space group $C2$, with unit-cell parameters $a = 85.98$, $b = 29.62$, $c = 40.33$ Å, $\beta = 101.3^\circ$. Crystals of SeMet-substituted N-DCLK (Leu120Met) were also obtained, but they exhibit the symmetry of space group $P2_1$, with unit-cell parameters $a = 38.81$, $b = 29.43$, $c = 40.1$ Å, $\beta = 115.7^\circ$.

1. Introduction

Neurons migrate long distances to form the complex laminar structures of the cerebral cortex. Neural precursors proliferate in the ventricular zone and post-mitotic neurons migrate to establish the mature layered cortex in waves, such that early-born neurons make up the deeper layers of the cortical cortex, whereas later-born neurons occupy the more superficial layers (Gupta *et al.*, 2002). In this way, a six-neuronal layered neocortex develops. There are several types of human cortical disorders which are associated with defects in neuronal migration. Two genes have been cloned that, when mutated, result in lissencephaly: lissencephaly 1 (*Lis1*) on chromosome 17p13.3 (Reiner *et al.*, 1993) and doublecortin (*DCX*) on Xq22.3-q24 (des Portes *et al.*, 1998; Gleeson *et al.*, 1998). The amino-acid sequence reveals that Lis1 is a member of the WD40 repeat superfamily (Reiner *et al.*, 1993, 1995), related structurally to the β-subunits of trimeric G-proteins. Lis1 appears to interact with microtubules and affects microtubule dynamics (Sapir *et al.*, 1997). The other protein, doublecortin (DCX), is 360 residues long and also associates with and stabilizes microtubules and affects microtubule dynamics (Horesh *et al.*, 1999; Gleeson *et al.*, 1999; Francis *et al.*, 1999). Mutations in the X-linked *DCX* gene result in lissencephaly in males or subcortial laminar heterotopia (SCLH) in females (Berg *et al.*, 1998; des Portes *et al.*, 1998; Gleeson *et al.*, 1998). The distribution of patient mutations in the gene and secondary-structure prediction suggest that DCX contains a tandem of two homologous domains (Taylor *et al.*, 2000; Sapir *et al.*, 2000). Interestingly, similar tandems are found in other human proteins: the retinitis pigmentosa 1 (RP1) protein (Sullivan *et al.*, 1999) and doublecortin-like kinase (DCLK; Burgess *et al.*, 1999; Burgess & Reiner, 2000; Lin *et al.*, 2000) (Fig. 1*a*). The role of RP1 is not clear, but its similarity to DCX in the N-terminal region suggests that it may function in development or maintenance of the neuronal retina or neuronal components of photoreceptor cells (Sullivan *et al.*, 1999). The 729-residue-long DCLK exhibits about 70% similarity to DCX in its N-terminus, whereas the C-terminal domain contains a putative Ca^{2+}/calmodulin-dependent kinase (Lin *et al.*, 2000). This protein has also been shown to have microtubule binding, polymerization and stabilization properties and to coexpress with DCX in migrating neurons (Burgess *et al.*, 1999; Lin *et al.*, 2000), implicating that the two proteins may be involved in the same signaling pathway that regulates microtubules in migrating neurons.

In order to better understand the function of DCX-like domains, we have initiated a structural study with the aim of enhancing our understanding of normal cortical development and the pathogenesis of brain malformation. Here, we report the purification of three different fragments of human DCLK and the successful crystallization and preliminary

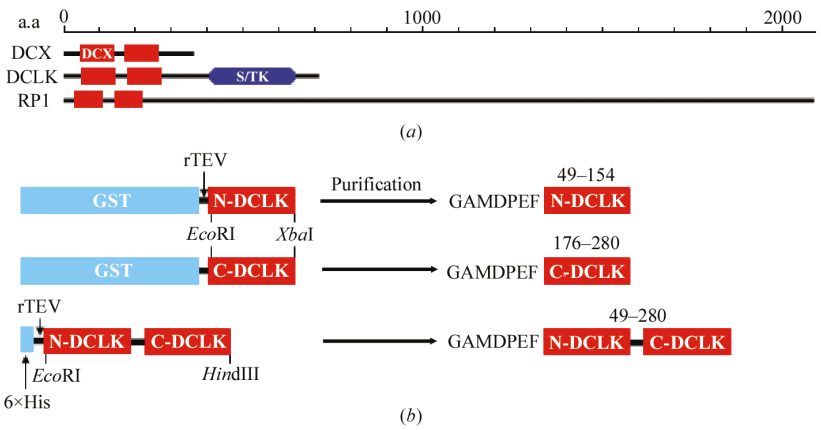

Figure 1
Domain architectures of human proteins containing DCX-like domains. (*a*) DCX, human doublecortin; DCLK, human doublecortin-like kinase; RP1, *Papio hamadryas* retinitis pigmentosa 1 protein (from RPS-BLAST 2.2.3). The red-coloured boxes indicate DCX-like domains and the S/TK describes serine/threonine protein kinase. (*b*) DCLK constructs used in this study.

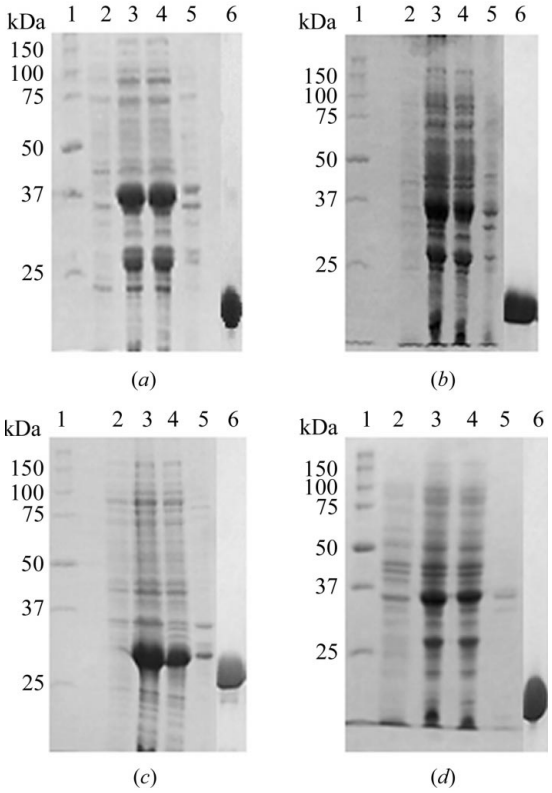

Figure 2
Expression and purification of DCLK constructs as judged by 12% SDS–PAGE. (*a*) N-DCLK, (*b*) C-DCLK, (*c*) tandem, (*d*) SeMet-N-DCLK(Leu120Met). Lane 1, protein-size markers; lane 2, uninduced *E. coli* BL21 (DE3) cell lysate; lane 3, induced cell lysate; lane 4, soluble fraction; lane 5, insoluble fraction; lane 6, purified protein.

X-ray data for the N-terminal domain (N-DCLK).

2. Protein expression and purification

Initially, we made three constructs: N-DCLK (residues 49–154), C-DCLK (residues 176–280) and a full tandem (residues 49–280) of DCLK (Fig. 1*b*). The DNA fragments coding for the respective fragments were amplified from the full-length cDNA, DCAMKL1, by polymerase chain reaction (PCR) using the primers 5′-CCG GAA TTC ACG CTC AGC TCC GAG-3′ and 5′-TGC TCT AGA TCA GAC GTT CAC CGA CC-3′; 5′-CCG GAA TTC GTG CGA GAG AAT AAG G-3′ and 5′-TGC TCT AGA TCA CTT TAC CAC TCG ACA TTC-3′; 5′-CCG GAA TTC ACG CTC AGC TCC GAG-3′ and 5′-CCC AAG CTT TCA CTT TAC CAC TCG ACA TTC-3′, respectively. The primers carried *Eco*RI and *Xba*I restriction sites for N-DCLK and C-DCLK, respectively, and *Eco*RI and *Hin*dIII sites for the tandem. The PCR products were gel-purified (Qiagen). N-DCLK and C-DCLK were subcloned into the *Eco*RI and the *Xba*I sites of the pGSTUni1 expression vector (Sheffield *et al.*, 1999) and the tandem was subcloned into the *Eco*RI and the *Hin*dIII sites of the pHisUni1 expression vector (Sheffield *et al.*, 1999). The resulting constructs, pGSTUni1-N-DCLK, pGSTUni1-C-DCLK and pHisUni1-tandem, were verified by restriction digestion and DNA sequencing. *Escherichia coli* strain BL21 (DE3) (Novagen) was transformed with the expression plasmids and selected on Luria–Bertani (LB) agar plates containing ampicillin. *E. coli* BL21 (DE3) cells harboring the plasmids were grown in LB medium (2.8 l) containing ampicillin at 310 K until an OD_{600} of between 0.6 and 0.8 was reached. After being cooled to 293 K, the cells were induced with 1 m*M* isopropyl-β-D-thiogalactopyranoside (IPTG) and grown for a further 14 h. The cells (∼15 g) were then harvested by centrifugation at 5000*g* for 20 min at 277 K. The cell pellets were resuspended in 100 ml ice-cold buffer *A* (50 m*M* Tris–HCl pH 7.5, 300 m*M* NaCl) and disrupted by ultrasonication. The crude cell extracts were centrifuged at 11 000*g* for 1 h at 277 K. The target proteins of the clear cell lysates were bound to 10 ml of glutathione Sepharose 4B (Pharmacia) equilibrated with buffer *A* on a Rocker II (Boekel Scientific) for 10 h at 277 K. After washing the resin with 1 l of buffer *A*, the bound glutathione S-transferase (GST) fusion proteins were eluted in 50 ml of buffer *A* containing 15 m*M* reduced glutathione. The eluted proteins were dialyzed against buffer containing 50 m*M* Tris–HCl pH 7.5, 1 m*M* DTT and 0.5 m*M* EDTA. GST was removed by incubating with rTEV protease (Gibco) followed by glutathione-affinity separation. Any residual GST was removed by size-exclusion chromatography (Superdex 75 column, Pharmacia). The cell lysate containing the His-tagged tandem was bound to 10 ml Ni–NTA agarose (Qiagen) equilibrated with buffer *A* for 10 h at 277 K. After the resin was washed with 1 l of buffer *A* containing 10 m*M* imidazole, the bound proteins were eluted in 50 ml of buffer *A* containing 200 m*M* imidazole. The His tag was released from the tandem by incubating with rTEV protease (Gibco), followed by Ni–NTA agarose chromatography and size-exclusion chromatography. After purification, all proteins had the same seven-residue cloning artifact (Gly-Ala-Met-Asp-Pro-Glu-Phe) at their N-termini (Fig. 1*b*). The final yield of the proteins was approximately 10 mg as determined by the Bradford procedure (Bio-Rad protein assay). The purified proteins were dialyzed against 20 m*M* Tris–HCl pH 7.0, concentrated to ∼20 mg ml^{-1} using a YM-3 membrane (Amicon) and stored at 193 K for use in crystallization trials. The homogeneity of the protein was assessed by 12% SDS–PAGE and Coomassie Blue staining (Fig. 2). The Leu120Met mutation was introduced into N-DCLK using QuikChange (Stratagene) to facilitate selenomethionine (SeMet) incorporation, as there are no naturally occurring methionines in the protein. In addition, six amino acids of the seven-residue cloning artifact (all except the Gly) were removed using QuikChange by the two-stage PCR

method (Wang & Malcolm, 1999) to improve the quality of the crystals. The mutated expression vector was transformed into the methionine auxotroph *E. coli* B834 (Novagen) and protein was expressed in minimal medium (2.8 l) supplemented with 50 mg ml^{-1} SeMet under the same conditions as the native plasmid. Purification of the SeMet-substituted protein was identical to that of native protein, with the exception that 5 m*M* dithiothreitol (DTT) was added to all buffers. The final yield of the purified protein was approximately 2 mg.

3. Crystallization

Crystallization of the purified proteins was initially carried out with Crystal Screen 1 and 2 from Hampton Research using the sitting-drop vapor-diffusion technique at 295 K. 1 µl of 10 mg ml^{-1} protein and 1 µl reservoir solution were mixed and equilibrated against reservoir solution. No crystals were obtained for the tandem or C-DCLK domain. Crystals of the N-DCLK domain were obtained directly from Crystal Screen 1 solution No. 47, containing 2.0 *M* $(NH_4)_2SO_4$ and 0.1 *M* sodium acetate trihydrate pH 4.6. A cluster of needle-like crystals (Fig. 3*a*) grew within 1 d. Optimization of the condition to 1.8 *M* $(NH_4)_2SO_4$, 0.1 *M* trisodium citrate pH 5.0 improved the quality of the crystals (Fig. 3*b*). However, single crystals suitable for high-resolution X-ray studies were only grown after seeding. Crystal setups containing 1.5 µl of 5 mg ml^{-1} protein solution and 1.5 µl reservoir solution containing 1.55 *M* $(NH_4)_2SO_4$ in 0.1 *M* trisodium citrate buffer pH 5.0 were prepared and equilibrated overnight against the reservoir solution. At the same time, a cluster of plate-like crystals was serially diluted 5, 25, 125 and 625 times in artificial mother liquor containing 1.8 *M* $(NH_4)_2SO_4$, 0.1 *M* trisodium citrate pH 5.0. 1 µl seed aliquots were placed in the equilibrated drops. Single plate-like crystals appeared within 1 d and grew for about three months to reach maximal dimensions of approximately 0.15 × 0.15 × 0.01 mm (Fig. 3*c*). The best results were obtained with 125-fold diluted seeds. SeMet-substituted protein crystals were also obtained from a slightly different crystallization condition using the sitting-drop vapour-diffusion technique at 295 K. Equal volumes of 15 mg ml^{-1} protein and reservoir solution containing 0.1 *M* trisodium citrate pH 5.2–5.8 and 2.2–2.4 *M* $(NH_4)_2SO_4$ were mixed and equilibrated against the reservoir solution. Plate-like clusters appeared within 1 d and took one to two weeks to grow into diffraction-quality crystals (Fig. 3*d*).

4. Data collection

The crystals were thin plates, no more than 10–20 µm thick, and special care was taken to freeze the crystals for data collection without damaging them. The native crystals were transferred to a cryoprotecting solution containing 1.7 *M* $(NH_4)_2SO_4$ and 30% glycerol in the same well in which they were grown. First, 10 µl of mother liquor was added to the well and 2.0 µl aliquots of cryoprotecting solution were then gradually added to the drop and the same amount removed from the drop after mixing. After complete exchange of the mother liquor with cryoprotecting solution, the crystals were fished out with a loop larger than their size to ensure that they did not bend and were flash-frozen by immersion in liquid nitrogen. The data were collected at SBC beamline 19ID at the Advanced Photon Source. The unit cell was autoindexed as *C*2, with unit-cell parameters $a = 85.98$, $b = 29.62$, $c = 40.33$ Å, $\beta = 101.3^\circ$. Assuming one molecule of N-DCLK (12.9 kDa) per asymmetric unit, the crystal volume per unit of protein mass is 1.95 Å^3 Da^{-1}, which corresponds to an approximate solvent content of 36.4% in the crystal. This dense packing probably accounts for the high-resolution data obtained in spite of the small size of the crystals. The SeMet crystals were frozen using 2.5 *M* $(NH_4)_2SO_4$ and 30% glycerol and multiwavelength anomalous dispersion (MAD) data (three wavelengths) were collected at beamline X9B of the NSLS (Brookhaven National Laboratory) to 1.6 Å

Table 1
Data-collection statistics.

Values in parentheses describe the relevant value for the last resolution shell.

		SeMet		
	Native	Inflection	Peak	Remote
Wavelength (Å)	0.9187	0.9795	0.9791	0.9718
Resolution range (Å)	1.5 (1.55–1.5)	1.6 (1.66–1.6)	1.6 (1.66–1.6)	1.6 (1.66–1.6)
Space group	*C*2	$P2_1$		
Unit-cell parameters (Å, °)	$a = 85.98$, $b = 29.62$, $c = 40.33$, $\beta = 101.3$	$a = 38.81$, $b = 29.43$, $c = 40.1$, $\beta = 115.7$		
Total reflections	47701	28275	28.259	29.160
Unique reflections	15460	9312	8997	9.217
Completeness (%)	96.8 (85.7)	84.6 (35.9)	81.9 (10.1)	83.7 (17.7)
$R_{\rm merge}$† (%)	8.4 (28.6)	3.8 (21.0)	4.2 (21.4)	3.2 (15.0)
$I/\sigma(I)$	14.1 (2.5)	32.5 (3.2)	31.0 (4.3)	41.7 (5.3)

† $R_{\rm merge} = \sum |I_i - \langle I \rangle| / \sum I$, where I_i is the intensity of the ith observation and $\langle I \rangle$ is the mean intensity of the reflections.

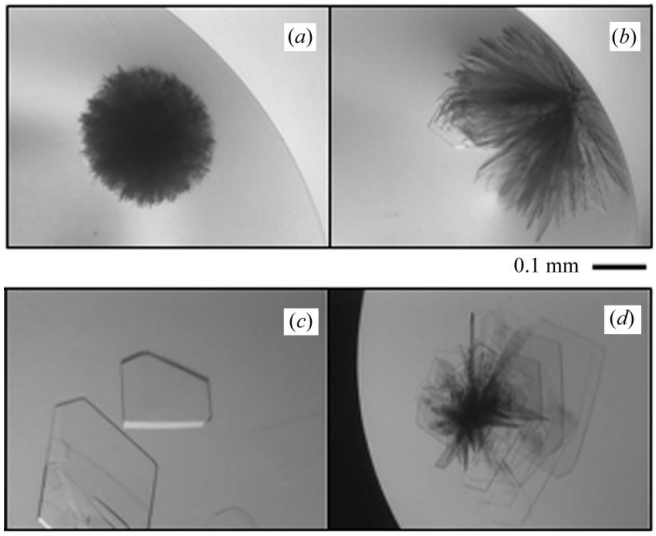

Figure 3
Photographs of N-DCLK crystals. (*a*) Needle-like cluster from 2.0 *M* $(NH_4)_2SO_4$, 0.1 *M* sodium acetate trihydrate pH 4.6. (*b*) Plate-like cluster from 1.8 *M* $(NH_4)_2SO_4$, 0.1 *M* trisodium citrate pH 5.0. (*c*) Single crystals grown by seeding of a pre-equilibrated drop containing 5 mg ml^{-1} protein solution and a reservoir solution consisting of 1.55 *M* $(NH_4)_2SO_4$, 0.1 *M* trisodium citrate buffer pH 5.0. (*d*) Plate-like cluster of SeMet-N-DCLK(Leu120Met) from 2.4 *M* $(NH_4)_2SO_4$, 0.1 *M* trisodium citrate pH 5.8.

resolution. The Se-labeled crystals belonged to space group $P2_1$, with unit-cell parameters a = 38.81, b = 29.43, c = 40.1 Å, β = 115.7°. Data were indexed, integrated and reduced using *HKL*2000 (Otwinowski & Minor, 1997) (Table 1). Structure determination will be reported elsewhere.

The work was funded by NIH grant NS36267 to ZSD. We thank Drs Wladek Minor and Andrzej Joachimiak for assistance with data collection at SBC beamline 19ID at the Advanced Photon Source.

References

Berg, M. J., Schifitto, G., Powers, J. M., Martinez-Capolino, C., Fong, C. T., Myers, G. J., Epstein, L. G. & Walsh, C. A. (1998). *Neurology*, **50**, 1143–1146.

Burgess, H. A., Martinez, S. & Reiner, O. (1999). *J. Neurosci. Res.* **58**, 567–575.

Burgess, H. A. & Reiner, O. (2000). *Mol. Cell. Neurosci.* **16**, 529–541.

Francis, F., Koulakoff, A., Boucher, D., Chafey, P., Schaar, B., Vinet, M. C., Friocourt, G., McDonnell, N., Reiner, O., Kahn, A., Denoulet, P., McConnell, S. K., Berwald-Netter, Y. & Chelly, J. (1999). *Neuron*, **23**, 247–256.

Gleeson, J. G., Allen, K. M., Fox, J. W., Lamperti, E. D., Berkovic, S., Scheffer, I., Cooper, E. C., Dobyns, W. B., Minnerath, S. R., Ross, M. E. & Walsh, C. A. (1998). *Cell*, **92**, 63–72.

Gleeson, J. G., Lin, P. T., Flanagan, L. A. & Walsh, C. A. (1999). *Neuron*, **23**, 257–271.

Gupta, A., Tsai, L. H. & Wynshaw-Boris, A. (2002). *Nature Rev. Genet.* **3**, 342–355.

Horesh, D., Sapir, T., Francis, F., Wolf, S. G., Caspi, M., Elbaum, M., Chelly, J. & Reiner, O. (1999). *Hum. Mol. Genet.* **8**, 1599–1610.

Lin, P. T., Gleeson, J. G., Corbo, J. C., Flanagan, L. & Walsh, C. A. (2000). *J. Neurosci.* **20**, 9152–9161.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Portes, V. des, Pinard, J. M., Billuart, P., Vinet, M. C., Koulakoff, A., Carrie, A., Gelot, A., Dupuis, E., Motte, J., Berwald-Netter, Y., Catala, M., Kahn, A., Beldjord, C. & Chelly, J. (1998). *Cell*, **92**, 51–61.

Reiner, O., Albrecht, U., Gordon, M., Chianese, K. A., Wong, C., Gal-Gerber, O., Sapir, T., Siracusa, L. D., Buchberg, A. M., Caskey, C. T. & Eichele, G. (1995). *J. Neurosci.* **15**, 3730–3738.

Reiner, O., Carrozzon, R., Shen, Y., Wehnert, M., Faustinella, F., Dobyns, W. B., Caskey, C. T. & Ledbetter, D. H. (1993). *Nature (London)*, **364**, 717–721.

Sapir, T., Elbaum, M. & Reiner, O. (1997). *EMBO J.* **16**, 6977–6984.

Sapir, T., Horesh, D., Caspi, M., Atlas, R., Burgess, H. A., Wolf, S. G., Francis, F., Chelly, J., Elbaum, M., Pietrokovski, S. & Reiner, O. (2000). *Hum. Mol. Genet.* **9**, 703–712.

Sheffield, P., Garrard, S. & Derewenda, Z. (1999). *Protein Expr. Purif.* **15**, 34–39.

Sullivan, L. S., Heckenlively, J. R., Bowne, S. J., Zuo, J., Hide, W. A., Gal, A., Denton, M., Inglehearn, C. F., Blanton, S. H. & Daiger, S. P. (1999). *Nature Genet.* **22**, 255–259.

Taylor, K. R., Holzer, A. K., Bazan, J. F., Walsh, C. A. & Gleeson, J. G. (2000). *J. Biol. Chem* **275**, 34442–34450.