Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Crystallization and preliminary studies of the DNA-binding runt domain of AML1

**Stefan Bäckström,[a] Sheng-Hua Huang,[a] Magnus Wolf-Watz,[b] Xiao-Qi Xie,[c] Torleif Härd,[b] Thomas Grundström[d] and Uwe H. Sauer[a]***

[a]Umeå Center for Molecular Pathogenesis, UCMP, Umeå University, SE-901 87, Umeå, Sweden, [b]Department of Biotechnology, Royal Institute of Technology, Stockholm, SE-141 57 Huddinge, Sweden, [c]CMM, Faculty of Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario K1G 8M5, Canada, and [d]Division of Tumour Biology, Department of Cell and Molecular Biology, Umeå University, SE-901 87 Umeå, Sweden

Correspondence e-mail: uwe@ucmp.umu.se



Received 26 June 2000
Accepted 31 October 2000

The acute myeloid leukaemia 1 (AML1) protein belongs to the Runx family of transcription factors and is crucial for haematopoietic development. The genes encoding Runx1 and its associated factor CBF$\beta$ are the most frequent targets for chromosomal rearrangements in acute human leukaemias. In addition, point mutations of Runx1 in acute leukaemias and in the familial platelet disorder FPD/AML cluster within the evolutionary conserved runt domain that binds both DNA and CBF$\beta$. Here, the crystallization of the Runx1 runt domain is reported. Crystals belong to space groups *C*2 and *R*32 and diffract to 1.7 and 2.0 Å resolution, respectively.

## 1. Introduction

Runx proteins, Runx1 to Runx3 (nomenclature according to HUGO will be used throughout this article), are the mammalian members of the runt family of heterodimeric transcription factors, named after the *Drosophila* gene runt.

Runx1, also termed AML1, CBF$\alpha$2 (core binding factor $\alpha$2) or PEBP2$\alpha$B (polyoma enhancer binding protein 2$\alpha$B), forms heterodimers with CBF$\beta$, also called PEBP2$\beta$, thus improving the DNA-binding affinity without involving CBF$\beta$–DNA contacts. Both proteins are essential for blood-cell formation. Homozygous disruption of either of the genes results in a total lack of definitive haematopoietic stem cells and consequently a lack of blood cells. The Runx1–CBF$\beta$ complex is also important in blood-vessel formation. Mice lacking a functional Runx1 gene die from massive haemorrhages. The second member of the runt domain protein family, Runx2 (also known as AML3, CBF$\alpha$1 and PEBP2$\alpha$A), is essential for osteoblast differentiation and bone formation, whereas the *in vivo* function of the third member, Runx3 (also known as AML2, CBF$\alpha$3 and PEBP2$\alpha$C), is unknown (Friedman, 1999; Namba *et al.*, 2000; Speck *et al.*, 1999; Suda *et al.*, 2000; Westendorf & Hiebert, 1999).

Taken together, the genes for Runx1 and CBF$\beta$ are the most frequent targets for chromosomal rearrangements in acute human leukaemias. All oncogenic variants of Runx1 contain the evolutionary conserved runt domain (Friedman, 1999; Speck *et al.*, 1999; Westendorf & Hiebert, 1999). Point mutations or deletions in the Runx1 gene are also found in several percent of acute leukaemias and in the familial platelet disorder FPD/AML and cluster within the runt domain (Lee *et al.*, 1997; Osato *et al.*, 1999; Song *et al.*, 1999; Westendorf & Hiebert, 1999).

Recently, the solution structure of the 128 amino-acid Runx1 runt domain was reported (Berardi *et al.*, 1999; Nagata *et al.*, 1999). The domain belongs to the S-type immunoglobulin fold family and is structurally related to the DNA-binding domains of STAT, NF-$\kappa$B, NFAT, p53 and the T-domain transcription factors.

In addition to binding CBF$\beta$, Runx1 interacts with and functions in conjunction with a number of proteins including several transcription factors, co-activators and co-repressors (Friedman, 1999; Ito, 1999; McLarren *et al.*, 2000; Speck *et al.*, 1999; Westendorf & Hiebert, 1999). Most of these interactions include the runt domain and/or regions from the N- or C-terminus that negatively regulate either its DNA or CBF$\beta$ binding. Detailed structural information will enable understanding of the molecular basis of the many runt protein functions and especially tumour-causing malfunctions. Here, we report the characterization of two crystal forms of the Runx1 runt domain diffracting to 1.7 and 2.0 Å resolution, respectively.

## 2. Materials and methods

### 2.1. Expression and purification

The mouse Runx1/AML1 runt domain construct comprises residues Ser46–Asp185, with Cys72 and Cys81 replaced by serine to improve protein solubility. The double mutation does not affect DNA binding. The protein was expressed in *Escherichia coli* BL21(DE3) cells and purified to homogeneity from inclusion bodies (Wolf-Watz *et al.*, 1999). The final product is a monomeric protein with a molecular weight of 15.5 kDa. Labelling with Se-Met for MAD experiments was performed in B834(DE3) cells (Novagen), a methionine auxotrophic strain of *E. coli*, following a stan-

**Table 1**
Crystallographic data-collection statistics.

Values for the last resolution shell are in parentheses.

| Space group | *C*2 | *R*32 |
|---|---|---|
| Temperature (K) | 100 | 100 |
| Unit-cell parameters (Å,°) | $a$ = 91.1<br>$b$ = 46.2<br>$c$ = 62.9<br>$\beta$ = 91.9 | $a = b$ = 110.7<br>$c$ = 117.3 |
| Total reflections | 111367 | 95246 |
| Unique reflections | 27534 | 17510 |
| Multiplicity | 4 | 5 |
| Resolution range (Å) | 26.0–1.7 (1.75–1.70) | 26.0–2.0 (2.09–2.02) |
| Completeness (%) | 97.2 (96.6) | 99.6 (99.9) |
| $I/\sigma(I) > 3$ (%) | 93.9 (89.0) | 90.5 (79.0) |
| $I/\sigma(I)$ | 13 (4.2) | 33 (2.1) |
| $R_{\rm merge}$†(%) | 4.1/15 | 4.5/17 |
| Mosaicity (°) | 0.52 | 0.73 |

† $R_{\rm merge}(I) = 100[\sum_{\mathbf{h}} \sum_{i=1}^{N} |I_i(\mathbf{h}) - \langle I(\mathbf{h})\rangle| / \sum_{\mathbf{h}} \sum_{i=1}^{N} I_i(\mathbf{h})]$.

dard protocol (Hendrickson *et al.*, 1990). Full Se-Met incorporation was confirmed by mass spectrometry. The purified protein was concentrated to 35 mg ml$^{-1}$ in storage buffer containing 140 m*M* NaCl and 100 m*M* $MgSO_4$, 20 m*M* hydroxyethyl piperazine ethane sulfonic acid (HEPES) at pH 8.0. Aliquots were immediately flash-cooled and stored in liquid nitrogen.

### 2.2. Crystallization

After thawing on ice the protein was diluted 1.5 times with buffer containing 1.6 *M* NaCl, 100 m*M* $MgSO_4$, 20 m*M* HEPES at pH 8.0 to a protein concentration of 23 mg ml$^{-1}$. An NaCl concentration of about 630 m*M* helped to keep the protein soluble. Initial crystallization trials using the hanging-drop method in combination with a sparse-matrix approach (Jancarik & Kim, 1991; Hampton Research) and an in-house screen resulted in crystals from a variety of PEG solutions at different pH values. To grow crystals in space group *C*2, 2 µl of the above solution containing native or Se-Met-labelled protein were premixed on a cover slip with 0.4–0.8 µl of a 14 base-pair DNA duplex. This resulted in molar DNA:protein ratios of 0.61–1.26. 2 µl of reservoir solution [25%(*w*/*v*) polyethylene-glycol (PEG) 3350, 16%(*v*/*v*) glycerol, 130 m*M* sodium cacodylate pH 6.4] was then added. The drops were sealed and equilibrated over 1 ml of reservoir solution. The blunt-ended DNA construct comprised the high-affinity (HA) Runx1 binding site, 5′-GCAAACCG-CAAACG-3′ (Thornell *et al.*, 1991). The DNA concentration was adjusted to 34 mg ml$^{-1}$ (MW = 8.52 kDa) in DNA buffer containing 200 m*M* NaCl, 50 m*M* $KH_2PO_4$ at pH 5.9. As a control, drops were set up in parallel with corresponding volumes (0.4–0.8 µl) of DNA buffer lacking DNA.

Crystals in space group *R*32 were produced by diluting the protein (35 mg ml$^{-1}$ in storage buffer) 2.2 times with 1.95 *M* NaCl, resulting in a protein concentration of 16 mg ml$^{-1}$ and an NaCl concentration of 1.1 *M*. On a cover slip, we added 2.5 µl of this pre-mix to 2 µl of reservoir solution [30%(*v*/*v*) monomethylether polyethylene glycol (MPEG) 350, 5–10%(*w*/*v*) PEG 3350 adjusted with 70 m*M* HEPES to pH 7.0 or with 70 m*M* sodium cacodylate to pH 6.5]. The drops were sealed and equilibrated over 1 ml of reservoir solution.

All crystals were grown by the vapour-diffusion method in hanging drops at 293 K (McPherson, 1982) using VDX plates (Hampton Research).

### 2.3. X-ray diffraction analysis

Preliminary X-ray diffraction analysis of native and heavy-atom soaked runt domain crystals was carried out at the X-ray crystallography beamline I711 at MAX-Lab synchrotron, Lund, Sweden. Data were collected on a MAR Research 345 imaging-plate detector with a crystal-to-detector distance of 229 mm, yielding a maximum resolution of 1.7 Å. Additional in-house data were collected on a DIP2030 double imaging-plate system (MAC Science) using Cu $K\alpha$ radiation of 1.54 Å produced by an FR-591 X-ray generator (Nonius). The beam was collimated to 0.5 mm. All crystals were mounted free-standing in nylon loops (Sauer & Ceska, 1997) and flash-cooled at 100 K in an open-flow nitrogen stream (Oxford Cryosystems Cryostream). Diffraction intensities were indexed, integrated and scaled with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) or *XDS* and *XSCALE* (Kabsch, 1988). Table 1 contains data-collection statistics and crystal parameters.

## 3. Results and discussion

Initially, the best crystals grew in 1–3 d at room temperature from conditions lacking DNA, equilibrated over PEG 3350 in the presence of $MgSO_4$ in the pH range 6–8. Multiple nucleation resulted in large amounts of microcrystals. The crystals diffracted to 2.6 Å with a high mosaic spread of 1.8°. The space group is *C*2, with unit-cell parameters $a$ = 91, $b$ = 46, $c$ = 63 Å, $\beta$ = 92°. Assuming two molecules to be present in the asymmetric unit, this corresponds to a $V_{\rm M}$ value of 2.13 Å$^3$ Da$^{-1}$ (Matthews, 1968).

The problems of poor diffraction and high mosaic spread were overcome by adding the DNA duplex to the protein solution prior to crystallization setups. This prevented microcrystals from appearing, reduced the crystal growth rate and resulted in fewer and larger crystals. Diamond-shaped crystals appeared after 1 d and grew in one week to full size (250 × 350 × 400 µm) (Fig. 1*a*). The resolution improved from 2.7 to 1.7 Å and the mosaic spread decreased from 1.8 to

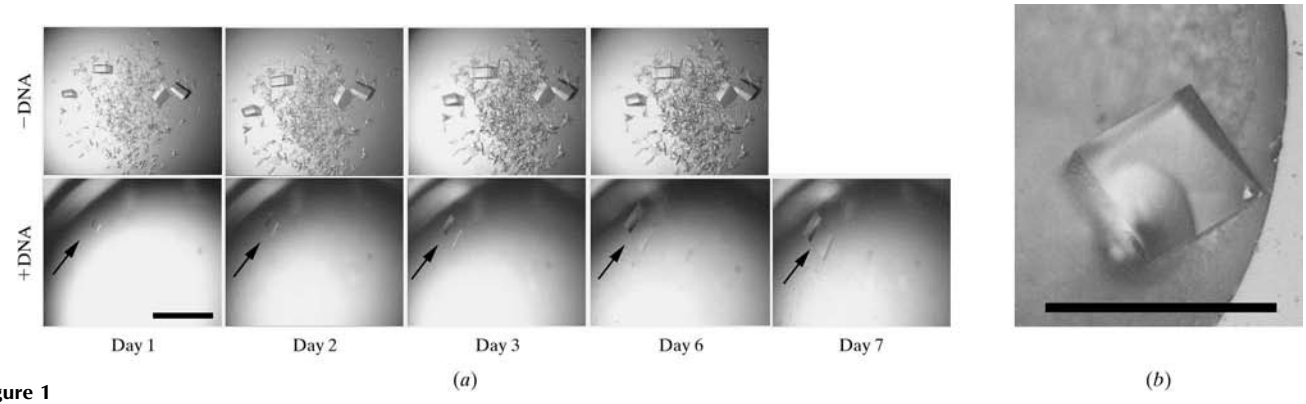


**Figure 1**
Two Runx1 runt domain crystal forms described in the text. (*a*) Crystals grown in space group *C*2. Top panel, crystals grown in the absence of DNA. Bottom panel, after addition of the high-affinity oligonucleotide the crystal growth rate is reduced and the size and quality improved. (*b*) Crystals in space group *R*32 were obtained in a salting-in experiment. (The bar corresponds to 1.0 mm in both cases.)

0.5°. These improvements were roughly proportional to the amount of DNA added up to a molar DNA:protein ratio of about 1:1. Further addition of DNA resulted in smaller crystals and at DNA:protein ratios much higher than 1:1 no crystals appeared at all. In control drops set up without DNA but with the corresponding volume of DNA buffer, there appeared first grainy precipitate and then small crystals grew interspersed with microcrystals within 2–3 d (Fig. 1*a*). These crystals were of inferior diffraction quality and resolution.

The space group and the unit-cell parameters of crystals grown in the presence of DNA did not change from the initial *C*2 crystals formed in the absence of DNA. This might indicate that the DNA is not incorporated into the crystal lattice. To analyse their contents, crystals were dissolved and the solution loaded onto an SDS–polyacrylamide gel. After separation and staining with ethidium bromide no trace of DNA could be detected (data not shown). It seems likely that some fraction of the protein binds to DNA, thus lowering the concentration of freely available protein in solution and reducing the nucleation rate of the apo runt domain. In order to investigate the crystallization conditions further, we added a DNA construct of same length and base composition and hence the same MW as the high-affinity binding site but with a randomized sequence. Crystals grew in 2–3 d from initial heavy precipitation and were of inferior quality, containing cracks. On the other hand, by lowering the protein concentration to values of 10 mg ml$^{-1}$ in the absence of DNA, we were able to slow down the rate of crystal growth and obtain improvements in crystal quality comparable with the crystals grown in the presence of the high-affinity DNA-binding site.

In order to avoid aggregation, it was crucial to keep the protein in high-salt solution (>600 m*M* NaCl, 100 m*M* $MgSO_4$). Therefore, the salting-in method in the absence of DNA was tried in hanging drops with a starting concentration of 1.1 *M* NaCl in the protein solution and no salt in the reservoir solution. This delayed the appearance of crystals from 1 d to one week. Crystals in the shape of cubes continued to grow for 1–2 months to final dimensions of $600 \times 600 \times 400$ µm in the new space group *R*32 (Fig. 1*b*). The unit-cell parameters are $a = b = 110.7$, $c = 117.3$ Å, resulting in a $V_M$ value of 4.46 Å$^3$ Da$^{-1}$, assuming one molecule in the asymmetric unit. Crystals diffract to 2.0 Å resolution with a mosaic spread of 0.7°.

Attempts to derivatize runt domain crystals with heavy metals failed. Therefore, Se-Met labelled protein was produced, which crystallized isomorphously in space group *C*2 in the presence of the high-affinity DNA. Diffraction tests using our in-house X-ray source resulted in diffraction to about 1.7 Å. The Runx1 runt domain structure was recently solved using three-wavelength multiple anomalous dispersion (MAD) data (Hendrickson, 1991) collected at the ESRF beamline ID14-EH4 and will be published elsewhere.

We thank Y. Cerenius for help with data collection at Max Lab, Lund, Sweden and S. McSweeney for help with MAD data collection at beamline ID14-EH4, ESRF, France.

## References

Berardi, M. J., Sun, C., Zehr, M., Abildgaard, F., Peng, J., Speck, N. A. & Bushweller, J. H. (1999). *Structure Fold. Des.* **7**, 1247–1256.
Friedman, A. D. (1999). *Leukemia*, **13**, 1932–1942.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Ito, Y. (1999). *Genes Cells*, **4**, 685–696.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Lee, B., Thirunavukkarasu, K., Zhou, L., Pastore, L., Baldini, A., Hecht, J., Geoffroy, V., Ducy, P. & Karsenty, G. (1997). *Nature Genet.* **16**, 307–310.
McLarren, K. W., Lo, R., Grbavec, D., Thirunavukkarasu, K., Karsenty, G. & Stifani, S. (2000). *J. Biol. Chem.* **275**, 530–538.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: John Wiley & Sons.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nagata, T., Gupta, V., Sorce, D., Kim, W. Y., Sali, A., Chait, B. T., Shigesada, K., Ito, Y. & Werner, M. H. (1999). *Nature Struct. Biol.* **6**, 615–619.
Namba, K., Abe, M., Saito, S., Satake, M., Ohmoto, T., Watanabe, T. & Sato, Y. (2000). *Oncogene*, **19**, 106–114.
Osato, M., Asou, N., Abdalla, E., Hoshino, K., Yamasaki, H., Okubo, T., Suzushima, H., Takatsuki, K., Kanno, T., Shigesada, K. & Ito, Y. (1999). *Blood*, **93**, 1817–1824.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sauer, U. H. & Ceska, T. A. (1997). *J. Appl. Cryst.* **30**, 71–72.
Song, W. J., Sullivan, M. G., Legare, R. D., Hutchings, S., Tan, X., Kufrin, D., Ratajczak, J., Resende, I. C., Haworth, C., Hock, R., Loh, M., Felix, C., Roy, D. C., Busque, L., Kurnit, D., Willman, C., Gewirtz, A. M., Speck, N. A., Bushweller, J. H., Li, F. P., Gardiner, K., Poncz, M., Maris, J. M. & Gilliland, D. G. (1999). *Nature Genet.* **23**, 166–175.
Speck, N. A., Stacy, T., Wang, Q., North, T., Gu, T. L., Miller, J., Binder, M. & Marin-Padilla, M. (1999). *Cancer Res.* **59**, 1789S–1793S.
Suda, T., Takakura, N. & Oike, Y. (2000). *Int. J. Hematol.* **71**, 99–107.
Thornell, A., Hallberg, B. & Grundström, T. (1991). *J. Virol.* **65**, 42–50.
Westendorf, J. J. & Hiebert, S. W. (1999). *J. Cell Biochem.* **32/33** Suppl., 51–58.
Wolf-Watz, M., Xie, X. Q., Holm, M., Grundström, T. & Härd, T. (1999). *Eur. J. Biochem.* **261**, 251–60.