# crystallization papers

# Identification, cDNA cloning, expression, crystallization and preliminary X-ray analysis of an exceptionally halotolerant carbonic anhydrase from *Dunaliella salina*

**Lakshmanane Premkumar,[a,b]
Harry M. Greenblatt,[b] Umesh K.
Bageshwar,[a]† Tatyana
Savchenko,[a] Irena Gokhman,[a]
Ada Zamir[a] and Joel L.
Sussman[b]***

[a]Department of Biological Chemistry,
Weizmann Institute of Science, Rehovot 76100,
Israel, and [b]Department of Structural Biology,
Weizmann Institute of Science, Rehovot 76100,
Israel

† Present address: Institute of Plant Genomics
and Biotechnology, Department of Horticulture,
Texas A&M University, College Station,
TX 77843, USA.

Correspondence e-mail:
joel.sussman@weizmann.ac.il

An extracellular $\alpha$-type carbonic anhydrase (dCAII) from the salt-tolerant alga *Dunaliella salina* differs from its mesophilic counterparts in remaining active from zero to multimolar salt concentrations. To gain insight into the outstanding salt tolerance of dCAII, the enzyme was functionally overexpressed in *Escherichia coli*, purified by affinity chromatography and crystallized by the hanging-drop method. The crystals belonged to space group $P2_1$, with unit-cell parameters $a = 47.0$, $b = 119.9$, $c = 58.5$ Å, $\beta = 94.2°$. Data from a single crystal were collected to 2.4 Å resolution under cryogenic conditions (120 K) using an R-AXIS IV$^{++}$ detector mounted on a Rigaku RU-H3R rotating-anode generator. The asymmetric unit contains two molecules of the protein, which corresponds to $V_M$ = 2.65 Å$^3$ Da$^{-1}$ and a solvent content of 52.7%.

## 1. Introduction

Proteins normally function within a narrow range of relatively low salt concentrations. In contrast, halophilic proteins derived from extreme halophilic archaea usually require multimolar salt concentrations for correct folding and activity (Madern *et al.*, 2000; Mevarech *et al.*, 2000).

Recently, members of another class of proteins, designated as halotolerant, were identified in the extremely salt-tolerant unicellular green alga *Dunaliella salina* (Gokhman *et al.*, 1999). *D. salina* cells grow in salinities ranging from <0.1 *M* to near saturation of NaCl, maintain a low intracellular ionic concentration and achieve osmotic balance by intracellular accumulation of glycerol (Avron, 1986). Hence, the intracellular proteins are not challenged by salt, while the extracellular proteins are expected to remain active under widely differing salinities, extending over the ranges suitable for mesophilic and halophilic proteins.

The first identified protein fulfilling this expectation was dCAI (originally named Dca), a 60 kDa protein containing two linked repeats of $\alpha$-type carbonic anhydrase sequences (P54212). dCAI exhibits enzymatic activity over the range from 0 to >2.0 *M* NaCl. In composition, dCAI differed from other $\alpha$-type carbonic anhydrases in containing an excess of acidic over basic amino-acid residues (Fisher *et al.*, 1996).

Recently, we identified in *D. salina* an additional $\alpha$-type carbonic anhydrase: the 30 kDa non-duplicated dCAII. This enzyme is ~57% identical to each N and C repeat of dCAI and is likewise extracellularly oriented, highly acidic and functionally salt tolerant (Premkumar *et al.*, 2003).

Towards the elucidation of the structural elements underlying the unusual salt responses of the *D. salina* carbonic anhydrases, we set out to determine the three-dimensional structure of dCAII. Here, we report the initial detection, cDNA cloning, heterologous expression, purification, functional characterization, crystallization and preliminary X-ray analysis of dCAII. Structural comparisons of dCAII with numerous structures determined for mesophilic carbonic anhydrases will help to elucidate the principles that underlie the high degree of salt tolerance of the *D. salina* enzyme.

## 2. Experimental and results

### 2.1. Detection and cDNA cloning of dCAII

Affinity chromatography on a column of agarose-coupled *p*-aminomethylbenzene sulfonamide (Sigma) (Whitney, 1974) of extracts of *D. salina* grown in 3.5 *M* NaCl (Fisher *et al.*, 1994; Sadka *et al.*, 1989) yielded a fraction eluted with 500 m*M* NaClO$_4$ at pH 5.6 that predominantly contained the 60 kDa dCAI (Bageshwar *et al.*, unpublished results), consistently accompanied by a minor ~30 kDa protein (dCAII). After its elution from an SDS–PAGE gel, dCAII was subjected to partial digestion with V8 protease and the N-terminal sequence of several peptides was determined as described in Fisher *et al.* (1996). A sequence of 48 amino-acid residues, compiled from two overlapping peptides, showed homology to $\alpha$-type carbonic anhydrases and served to design forward and
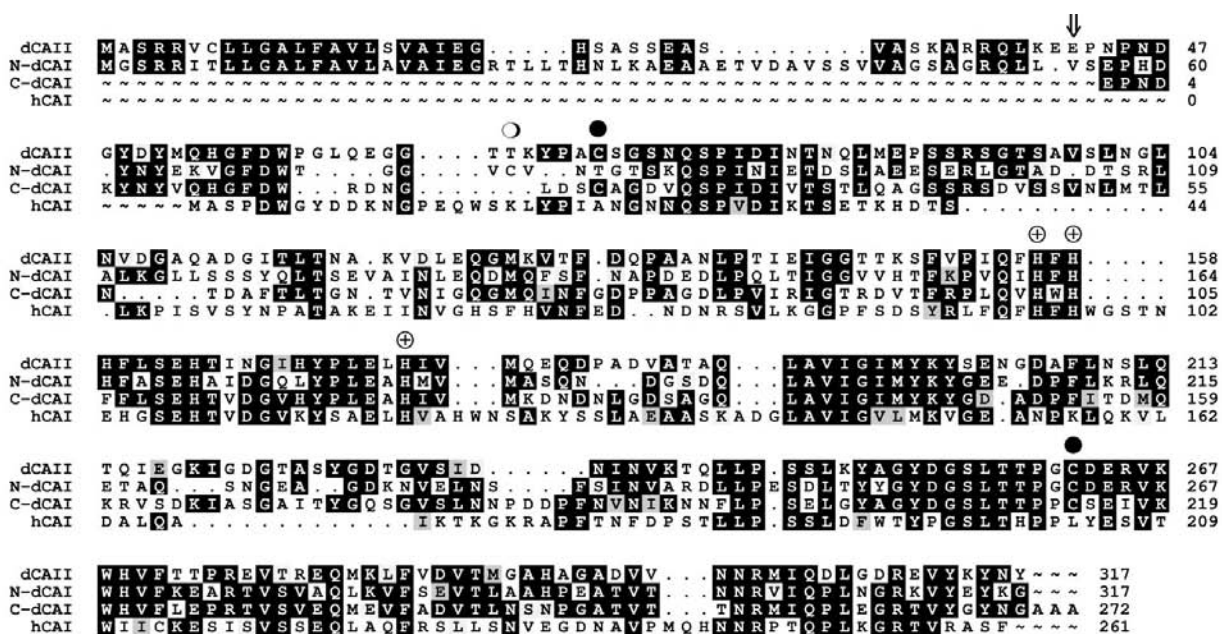
**Figure 1**

The complete amino-acid sequence of dCAII aligned with N and C repeats of dCAI (P54212) and hCAI (human carbonic anhydrase isozyme I; P00915). The arrow shows the N-terminus of mature dCAI after cleavage of the leader peptide (Fisher *et al.*, 1996) and homology-based designation of the N-terminus of mature dCAII. The circled plus signs label conserved His residues coordinating with Zn at the active site, filled circles label conserved Cys residues participating in disulfide-bond formation in extracellular carbonic anhydrases and the empty circle labels the Cys75 residue in N-Dca that is probably involved in disulfide-bond formation with Cys262.

reverse degenerate primers. These were used in RT-PCR (reverse transcriptase polymerase chain reaction) together with total RNA isolated from cells grown in 3.5 *M* NaCl as a template to amplify a 131 bp cDNA sequence. 3′ and 5′ RACE (rapid amplification of cDNA sequence ends) procedures were subsequently employed to obtain the 1350 bp full-length cDNA that encoded a protein of 317 amino-acid residues (dCAII).

Sequence alignments (Fig. 1) verified the identity of dCAII as an α-type carbonic anhydrase preceded by a leader peptide of 41 amino acids, of which 27 showed identity to the N-terminal leader peptide of dCAI (Fisher *et al.*, 1996). Based on this alignment, Glu41 in dCAII was assigned as the N-terminal residue of the mature protein (Fig. 1).

## 2.2. Plasmid construction, heterologous expression and purification

The cDNA for mature dCAII (residues 41–317) was synthesized by PCR (30 cycles at 368 K for 1 min, 327 K for 2 min and 345 K for 3 min) with the full-length cDNA clone as template and the primers 5′-GGA-**AGGATC**CGAAGAACCCAACCCCAA-TG-3′ (forward) and 5′-ATTG**CTCGAG**-TTAATAGTTGTACTTGTACAC-3′ (reverse). The primers were designed to include *Bam*HI and *Xho*I restriction sites (in bold), respectively. The PCR product was digested with *Bam*HI and *Xho*I and the 843 bp product was inserted into the pET21a vector (Novagen), pre-digested with *Bam*HI and *Xho*I, to create the expression vector pdCAII. The insert was fused in-frame to upstream vector sequences containing a translation-initiation site followed by a sequence encoding a T7 tag (see Novagen catalogue for sequences). The recombinant plasmid was transformed into the *E. coli*
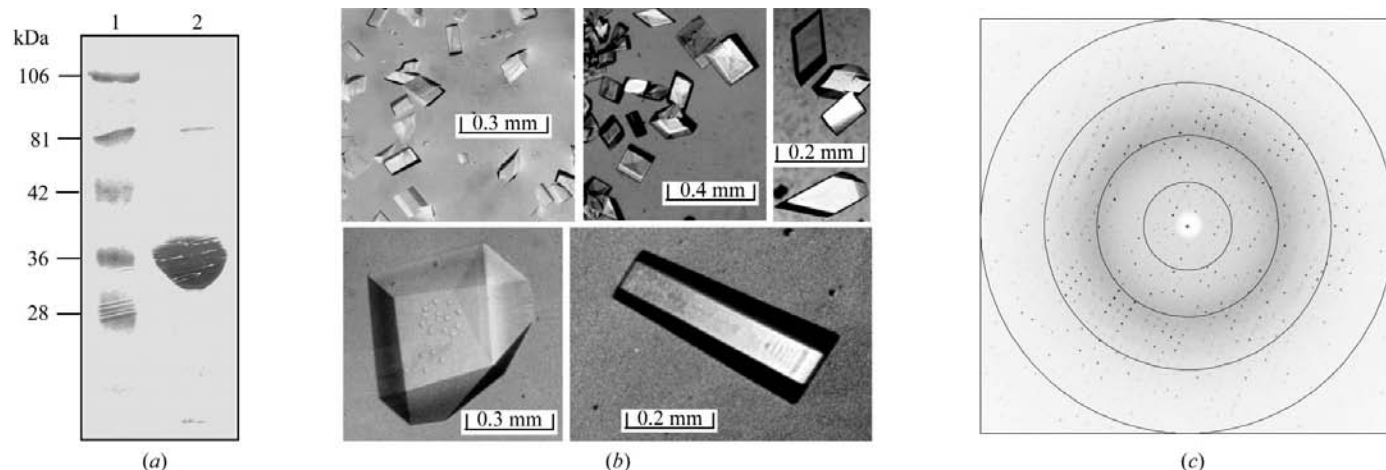


**Figure 2**

(*a*) Affinity-purified dCAII analyzed on SDS–PAGE (lane 2), size markers (lane 1). (*b*) Primitive monoclinic dCAI crystals grew in 2-propanol, PEG 4000 and trisodium citrate pH 5.6 within 2–3 d and ranged from 0.1 to 0.8 mm in size. (*c*) Diffraction pattern extending beyond 2.4 Å of dCAII crystals. The rings from the centre correspond to 9.77, 4.88, 3.24 and 2.44 Å.

# crystallization papers

Origami B strain B(DE3)pLysS [(F⁻ *ompT* *hsdS_B* (r_B⁻m_B⁻) *lacY1 gor522*::T*n*10 (Tc^R) *trxB*::kan (DE3) pLysS (Cm^R)]. Large-scale expression of dCAII was carried out in a 12 l culture grown in a fermentor, essentially as described for dCAI (Premkumar *et al.*, 2003).

Affinity purification of recombinant dCAII was carried out essentially as described for recombinant dCAI (Premkumar *et al.*, 2003). Specifically, 10 g of pelleted *E. coli* cells enriched in dCAII was suspended in 40 ml lysis buffer and lysed by repeated freezing in liquid nitrogen and thawing in an ice-water bath as described in Premkumar *et al.* (2003). The lysate was spun at 42 000 rev min⁻¹ for 1 h and the supernatant was mixed with a 10 ml packed volume of agarose-bound *p*-aminomethylbenzene sulfonamide pre-equilibrated in 25 m$M$ Na$_2$SO$_4$ in 25 m$M$ Tris–SO$_4$ pH 8.8; the suspension was shaken for 4 h at 277 K and poured into a column. The bulk of the non-specifically bound proteins were removed by washing with 300 ml equilibration buffer followed by 300 ml 100 m$M$ NaClO$_4$, 50 m$M$ Na$_2$SO$_4$, 25 m$M$ Tris–SO$_4$ pH 8.8 (*cf.* Whitney, 1974). The enzyme was eluted with 200 m$M$ NaClO$_4$, 100 m$M$ sodium acetate pH 5.6 as described in Premkumar *et al.* (2003). Complete EDTA-free protease-inhibitor cocktail (Roche Molecular Biochemicals; one tablet per 50 ml buffer solution) was added to all solutions used. The eluted fraction was extensively dialyzed against 25 m$M$ Tris–SO$_4$ pH 8.8, 10 μ$M$ ZnSO$_4$ and concentrated to ∼12.5 mg ml⁻¹ in an Amicon Centriprep YM-30. All purification steps were carried out at 277 K. SDS–PAGE analysis of the purified protein is shown in Fig. 2(*a*).

The recombinant dCAII was assayed for CO$_2$ hydration and esterase activities in increasing salinities. The results indicated that both activities were retained from zero to multimolar concentrations of NaCl (data not shown). Similar results had previously been observed for dCAI (Fisher *et al.*, 1996; Bageshwar *et al.*, unpublished results).

## 2.3. Crystallization

Crystallization was performed at 292 K using the hanging-drop vapour-diffusion method. The crystallization conditions were determined by a sparse-matrix screen using Crystal Screen and Crystal Screen II kits (Hampton Research). For the initial screen, drops were prepared by mixing 1 μl of protein solution (12.5 mg ml⁻¹ in 10 μ$M$ ZnSO$_4$, 25 m$M$ Tris–SO$_4$ pH 8.8) with 1 μl

**Table 1**
Diffraction data.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| Crystal system | Monoclinic |
| Space group | $P2_1$ |
| Crystal-to-detector distance (mm) | 200.2 |
| Data sweep (°) | 126 |
| Unit-cell parameters (Å, °) | $a = 47.0$, $b = 119.9$, $c = 58.5$, $\alpha = 90$, $\beta = 94.2$, $\gamma = 90$ |
| Matthews coefficient (Å³ Da⁻¹) | 2.65 |
| Unit-cell volume (Å³) | 328780 |
| No. of molecules per unit cell | 4 |
| No. of molecules in the AU | 2 |
| Total No. of unique reflections collected | 24493 |
| Overall redundancy | 2.56 |
| No. of observed reflections | 165404 |
| Wavelength used (Å) | 1.5418 |
| Resolution range (Å) | 35.0–2.4 (2.49–2.4) |
| Completeness (%) | 96.1 (85.8) |
| $R_{merge}$ | 0.026 (0.036) |
| $\langle I/\sigma(I) \rangle$ | 46.6 |
| Reflections <3$\sigma$ in highest resolution shell | 2.8 |
| Mosaicity (°) | 0.47 |

precipitant solution. The reservoir precipitant volume was 0.5 ml.

A precipitate formed in Screen 41 (Crystal Screen), containing 20% 2-propanol, 20% PEG 4000, 0.1 $M$ trisodium citrate pH 5.6, was the starting point for optimization trials. Dilution of the precipitant or lowering the protein concentration produced clear drops or precipitates, but no crystals. However, a variation on this screen produced many beautiful albeit small crystals (<0.1 mm) after 2–3 d by using the same precipitant concentration in the reservoir but preparing the drop with 1 μl protein solution together with 1 μl of precipitant which had been diluted 1:6 with water. When the total drop size was increased to 6 μl, retaining a 1:1 ratio of protein to diluted precipitant (1:6), fewer but significantly larger crystals (0.1–0.7 mm) varying in size and shapes (chunky, diamond, plates, rod) were seen after 2–3 d (Fig. 2*b*).

## 2.4. Data collection

A native data set from a single chunky crystal (∼0.2 mm in each direction) was collected at 120 K with an R-AXIS IV detector on an RU-300 rotating-anode generator operating at 50 kV and 90 mA. Each frame was recorded with a 10 min exposure and 0.5° oscillation around $\varphi$. The crystals were very sensitive to drop dehydration and layering oil onto the drop containing crystals before manipulation was crucial. The crystals were transferred to oil and placed directly into a nitrogen stream for data collection. The crystal-to-detector distance was set to 200 mm so that the spots were well resolved. The data set was inte-

grated and scaled using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

The recombinant dCAII crystallized in space group $P2_1$, with unit-cell parameters $a = 47.0$, $b = 119.9$, $c = 58.5$ Å, $\beta = 94.2°$. A complete native data set for dCAII was collected to a resolution of 2.4 Å (Fig. 2*c*) and the statistics of the data processing are shown in Table 1. Assuming that two 30 kDa molecules are present in the asymmetric unit and given a space group of $P2_1$, the Matthews coefficient ($V_M$) was calculated to be 2.65 Å³ Da⁻¹, with an estimated solvent content of 52.7%.

Human carbonic anhydrases for which crystal structures have been determined (Stams & Christianson, 2000) share considerable sequence identity with dCAII (37% with the human carbonic anhydrase I). However, using the known coordinates for molecular replacement or treating the template as polyalanine did not enable solution of the structure of dCAII. Hence, we plan to use the anomalous scattering by the intrinsic catalytic Zn atom in dCAII to carry out a MAD experiment to determine the crystal structure of the protein. The structure will be used in comparisons with mesophilic counterparts in order to discern for the first time the distinctive structural features conferring salt tolerance on an α-type carbonic anhydrase.

## References

Avron, M. (1986). *Trends Biochem.* **11**, 5–6.
Fisher, M., Gokhman, I., Pick, U. & Zamir, A. (1996). *J. Biol. Chem.* **271**, 17718–17723.
Fisher, M., Pick, U. & Zamir, A. (1994). *Plant Physiol.* **106**, 1359–1365.
Gokhman, I., Fisher, M., Pick, U. & Zamir, A. (1999). *Microbiology and Biogeochemistry of Hypersaline Environments*, edited by A. Oren, pp. 203–213. Boca Raton, FL, USA: CRC Press.
Madern, D., Ebel, C. & Zaccai, G. (2000). *Extremophiles*, **4**, 91–98.
Mevarech, M., Frolow, F. & Gloss, L. M. (2000). *Biophys. Chem.* **86**, 155–164.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Premkumar, L., Bageshwar, U., Gokhman, I., Zamir, A. & Sussman, J. L. (2003). *Protein. Expr. Purif.* **28**, 151–157.
Sadka, A., Lers, A., Zamir, A. & Avron, M. (1989). *FEBS Lett.* **244**, 93–98.
Stams, T. & Christianson, D. W. (2000). *EXS*, pp. 159–174.
Whitney, P. (1974). *Anal. Biochem.* **57**, 467–476.