

Article

Structural and Biochemical Characterization of an Atypical α -Carbonic Anhydrase from the Tardigrade *Ramazzottius varieornatus*

Byung Hoon Jo ^{1,2,3} 

¹ Division of Applied Life Science (BK21 Four), Gyeongsang National University, Jinju 52828, Republic of Korea; jobh@gnu.ac.kr

² Anti-Aging Bio Cell Factory Regional Leading Research Center (ABC-RLRC), Gyeongsang National University, Jinju 52828, Republic of Korea

³ Division of Life Science and Research Institute of Life Science, Gyeongsang National University, Jinju 52828, Republic of Korea

Abstract

The tardigrade *Ramazzottius varieornatus* exhibits extraordinary resilience to extreme environmental stresses, yet the functional diversity of its proteome remains largely unexplored. In this study, the structural and biochemical characterization of RvCA5, an atypical α -carbonic anhydrase (CA) identified in *R. varieornatus*, is presented. Expression analysis in *E. coli* revealed the spontaneous formation of a truncated RvCA5 species, which was confirmed to be unrelated to signal peptide cleavage. RvCA5 exhibited distinct structural features, including extended intrinsically disordered regions (IDRs) at both termini. Unlike canonical α -CAs, RvCA5 exhibited negligible CO₂ hydration activity, which was partially enhanced by the removal of the N-terminal IDR, suggesting that this region acts as a dynamic entropic barrier hindering substrate diffusion. RvCA5 possesses multiple surface-exposed reactive cysteine residues, resembling the redox-sensing human CA 3. Notably, consistent with a predicted nuclear localization signal, in silico modeling predicted that RvCA5 can bind DNA via a positively charged patch near the C-terminal IDR. The DNA-binding capability of RvCA5 was experimentally demonstrated by electrophoretic mobility shift assays. Collectively, these findings suggest that RvCA5 potentially functions as a redox-responsive transcriptional regulator.

Keywords: tardigrade; *Ramazzottius varieornatus*; carbonic anhydrase; intrinsically disordered region; surface-exposed cysteine



Academic Editor: Chojiro Kojima

Received: 20 December 2025

Revised: 29 January 2026

Accepted: 30 January 2026

Published: 3 February 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The tardigrade, also known as the water bear, is a microscopic animal renowned for its extraordinary ability to withstand extreme physical and chemical stresses, including extreme temperatures, high pressures, radiation, and dehydration. *Ramazzottius varieornatus* is one of the tardigrade species whose genome has been sequenced [1]. Despite available genomic information, research on *R. varieornatus* proteins has been primarily limited to tardigrade-specific intrinsically disordered proteins (TDPs). Dsup (Damage Suppressor) is the most extensively studied TDP; it not only protects the tardigrade's DNA against irradiation and subsequent oxidative stress but also confers similar radiotolerance to other organisms upon heterologous expression [1–5]. Other TDPs include heat-soluble proteins known to increase desiccation resistance [6,7]. Although recent studies have investigated

a limited number of enzymes, such as metabolite phosphatase [8], DNA polymerase [9], mevalonate kinase [10], and superoxide dismutase [11], further research on a broader range of proteins is required to better understand the extraordinary capabilities of *R. varieornatus*.

Carbonic anhydrase (CA, EC 4.2.1.1) is a zinc-containing metalloenzyme that catalyzes the reversible hydration of carbon dioxide (CO_2): $\text{CO}_2 + \text{H}_2\text{O} \leftrightarrow \text{HCO}_3^- + \text{H}^+$ [12]. The enzyme-catalyzed reaction is one of the fastest known in biology, with a catalytic rate constant (k_{cat}) of up to $4.4 \times 10^6 \text{ s}^{-1}$ [13]. CA is categorized into eight classes (α , β , γ , δ , ζ , η , θ , and ι) based on discovery order; these classes have evolved via convergent evolution and share almost no sequence homology. CA plays critical physiological roles, including CO_2 transport and metabolism, biocalcification, pH homeostasis, and photosynthesis [14]. As an essential enzyme ubiquitous across virtually all living organisms, the study of CA is crucial for a broader biological understanding of the host organism. Moreover, CA has gained significant attention as a green biocatalyst for CO_2 capture and utilization [15]. For this application, CAs from extremophiles or extremotolerant organisms are highly desirable due to their stability under harsh physical (e.g., high temperature) and chemical (e.g., organic solvents) conditions [16]. In this context, the investigation of CA from the extremotolerant *R. varieornatus* would not only broaden fundamental biological insights into this essential enzyme but also provide promising candidate biocatalysts for industrial applications.

Genomic analysis of *R. varieornatus* predicted the existence of ten α -CA (RvCA) isoforms [1]. Previously, six of these isoforms were recombinantly expressed in *Escherichia coli*; however, only RvCA5 could be obtained in sufficient quantities and successfully purified via His₆-tag affinity chromatography [17]. Herein, the characterization of RvCA5 is reported using a combination of in silico structural prediction and in vitro biochemical assays. RvCA5 showed atypical characteristics, raising intriguing insight about its physiological function.

2. Results and Discussion

2.1. Sequence Characteristics of RvCA5

Sequence alignment of RvCA5 with selected α -CAs of mammalian and bacterial origins is shown in Figure 1. RvCA5 consists of 347 amino acids and has a theoretical molecular weight of 39.1 kDa (Table 1). The conserved residues, including the three Zn^{2+} -coordinating His residues and the proton shuttle His residue, suggest that RvCA is an α -CA [18]. The RvCA5 sequence showed N-terminal and C-terminal extensions not found in the other α -CAs. Any putative signal peptide was not predicted by SignalP6.0.

Table 1. Protein parameters of RvCAs.

Name (RvY ID, UniProt ID)	Protein Length	Molecular Weight (kDa)	No. Cysteine Residues ^a		pI
			Conserved	Non-Conserved (Surface-Exposed)	
RvCA1 (RvY_03263, A0A1D1UN98)	323	36.6	2	0 (0)	9.27
RvCA2 (RvY_01739, A0A1D1UL76)	354	39.5	2	0 (0)	5.91
RvCA3 (RvY_01195, A0A1D1UQT7)	414	46.4	2	1 (1)	8.62
RvCA4 (RvY_04176, A0A1D1UXK6)	364	41.4	2	1 (1)	6.25
RvCA5 (RvY_07018, A0A1D1V0M4)	347	39.1	0	6 (4)	6.85
RvCA6 (RvY_06685, A0A1D1UZF6)	347	38.7	2	0 (0)	6.17
RvCA7 (RvY_13508, A0A1D1VTD1)	333	38.3	2	2 (2)	8.30
RvCA8 (RvY_17267, A0A1D1W3U6)	383	42.5	2	4 (4)	7.11
RvCA9 (RvY_09162, A0A1D1V8H4)	325	37.3	2	1 (1)	8.16
RvCA10 (RvY_13109, A0A1D1VUC0)	344	39.0	2	2 (2)	5.90

^a Signal peptides were excluded from the calculation.

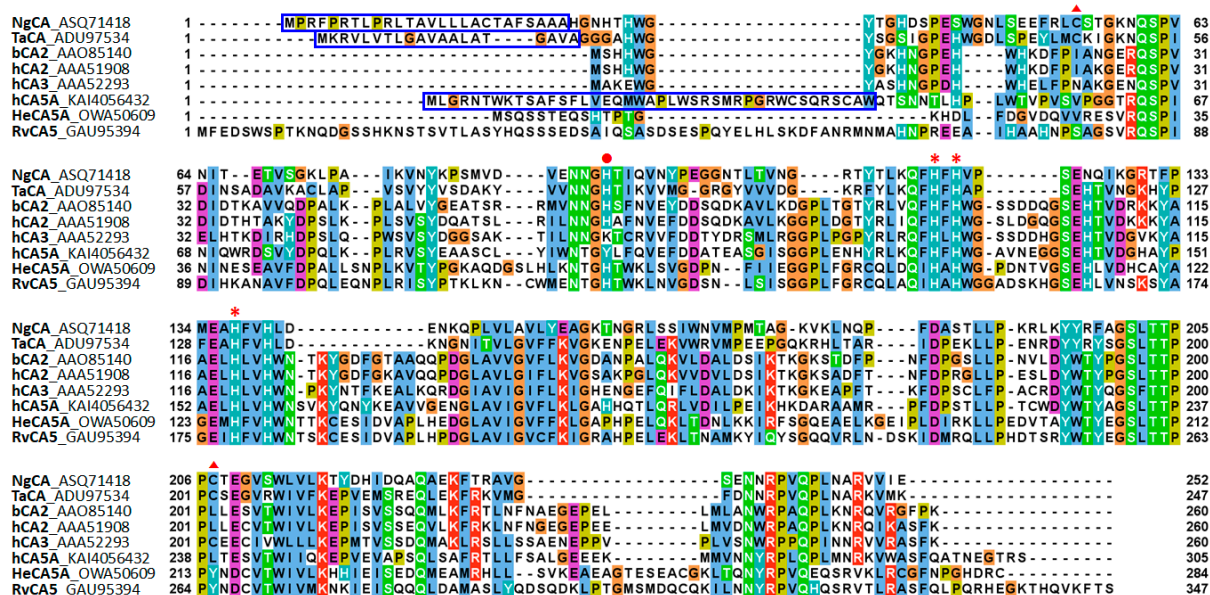


Figure 1. Multiple sequence alignments of selected α -CAs with RvCA5. The predicted signal peptide and the mitochondrial targeting sequence are boxed in blue. The two conserved cysteine residues for the formation of the intramolecular disulfide bond are indicated by a triangle (▲). The three zinc ligand histidine residues (*) and the proton shuttle residue (●) are marked.

A BlastP search using the RvCA5 sequence identified two tardigrade homologs: a putative mitochondrial carbonic anhydrase 5A (HeCA5A; GenBank: OWA50609) from *Hypsibius exemplaris* and a carbonic anhydrase 1-like protein (PmCA; GenBank: XP_055357851) from *Paramacrobiotus metropolitanus*, with sequence identities of 62% and 51%, respectively (Supplementary Figure S1a) [19]. When predicted using TargetP 2.0, DeepMito, and Euk-mPLoc, no putative mitochondrial targeting sequence (MTS) was found in RvCA5, HeCA5A, or PmCA, while an MTS was predicted in the sequence of human CA 5A (hCA5A), one of the 16 human α -CA isoforms whose subcellular localization has been experimentally verified [20]. Based on these results, it seems that these tardigrade CAs might not be localized in the mitochondria. Intriguingly, DeepLoc 2.1 identified a nuclear localization signal (NLS) in both the RvCA5 and HeCA5A sequences, suggesting their potential nuclear localization, whereas no NLS was found in the PmCA sequence.

In addition, transcriptomic data indicate that the RvCA5 gene is the most abundantly expressed among the ten RvCA genes (Table 2) [1], suggesting a dominant physiological role in the tardigrade. The mRNA levels of the RvCA5 gene are not significantly altered by dehydration stress and subsequent rehydration process (Table 2), suggesting that RvCA5 is constitutively expressed in the tardigrade.

Table 2. Expression profile of RvCA genes obtained by mRNA sequencing [1].

Name (RvY ID, UniProt ID)	Expression (FPKM)		
	Active	Dehydrated	Rehydration (3 h)
RvCA1 (RvY_03263, A0A1D1UN98)	5.7	3.3	3.0
RvCA2 (RvY_01739, A0A1D1UL76)	23.9	17.3	19.8
RvCA3 (RvY_01195, A0A1D1UQT7)	48.0	41.8	43.5
RvCA4 (RvY_04176, A0A1D1UXK6)	2.3	2.8	2.9
RvCA5 (RvY_07018, A0A1D1V0M4)	276.3	256.6	273.3
RvCA6 (RvY_06685, A0A1D1UZF6)	32.4	28.9	32.0
RvCA7 (RvY_13508, A0A1D1VTD1)	14.7	10.4	13.4
RvCA8 (RvY_17267, A0A1D1W3U6)	23.4	23.7	24.7
RvCA9 (RvY_09162, A0A1D1V8H4)	4.5	4.0	5.2
RvCA10 (RvY_13109, A0A1D1VUC0)	19.3	16.9	16.4

2.2. Expression and Purification of RvCA5 in *E. coli*

The RvCA5 gene was chemically synthesized with codon optimization for *E. coli*, cloned into the pET-22b(+) vector, and expressed in *E. coli* BL21(DE3) under the control of the strong T7lac promoter induced by the addition of 0.1 mM IPTG at 25 °C, as previously described [17]. The RvCA5 was highly overexpressed, appearing as thick bands on the SDS-PAGE gel with the expected molecular mass (40.1 kDa) (Figure 2a). After cell disruption by sonication, approximately half of the enzyme was found in the soluble fraction from which the RvCA5 was purified using His₆-tag affinity chromatography. During dialysis against a low-salt phosphate buffer, the precipitation of the RvCA5 occurred, and only a fraction of the enzyme was recovered in the supernatant after low-speed centrifugation (Figure 2a).

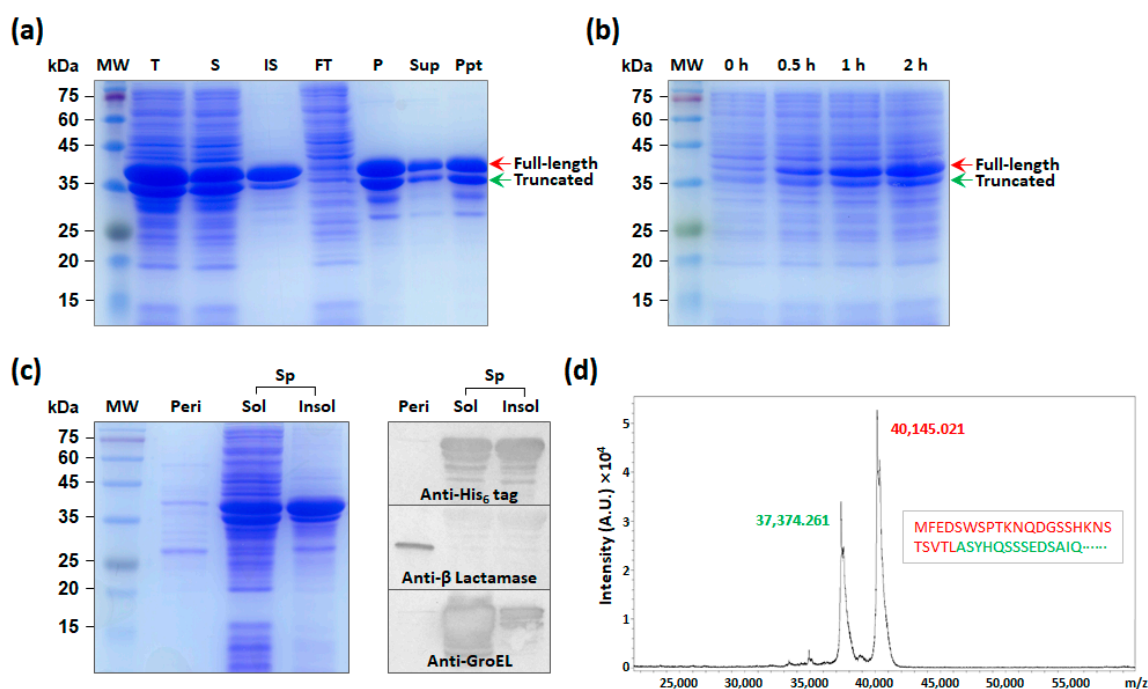


Figure 2. Analysis of recombinant expression, subcellular localization, and truncation of RvCA5 in *E. coli*. (a) Expression and purification of RvCA5 analyzed by SDS-PAGE followed by Coomassie Blue staining. Lanes: MW, molecular mass marker; T, total cell lysate; S, soluble fraction; IS, insoluble fraction; FT, flow-through; P, purified enzyme; Sup, supernatant; Ppt, precipitates. (b) Time-course analysis of RvCA5 expression. After cell harvest, cells were directly mixed with 4× Laemmli sample buffer and heated for sample preparation. (c) Subcellular localization of RvCA5. Lanes: MW, molecular mass marker; Peri, periplasmic fraction; Sp, spherolast fraction; S, soluble fraction; IS, insoluble fraction. (d) Mass spectrometric analysis of purified RvCA5.

2.3. Analysis of Truncated RvCA5

Notably, the RvCA5 appeared in two distinct bands on the SDS-PAGE gel (Figure 2a). The small-sized RvCA5 was identified as an N-terminal truncated form of RvCA5, since both proteins, corresponding to the two bands, were purified via the C-terminal His₆-tag (Figure 2a). The truncated RvCA5 was unlikely to have formed due to internal translation initiation, as no possible translation initiation site with a sufficiently high expression level and the proper reading frame was predicted using the UTR Designer. In addition, the truncation was not due to harsh cell disruption, and the truncated RvCA5 was generated early in the protein expression process and accumulated over time (Figure 2b).

It was hypothesized that the truncated RvCA5 was formed via the rapid but incomplete proteolytic cleavage of the N-terminus shortly after the full-length protein was made. One possibility is that the N-terminal region serves as a signal peptide, directing the entire

protein to be exported out of the cytoplasm into the periplasm, where it is then cleaved by a signal peptidase [21]. To confirm this possibility, the periplasmic fraction of the RvCA5-expressing cells was separated from the remaining spheroplast fraction, and then it was determined which fraction contained the RvCA5 proteins. The result showed that the His₆-tagged RvCA5 was only found in the spheroplast fraction and was absent in the periplasmic fraction (Figure 2c). The Western blot results using the periplasmic marker β -Lactamase and the cytoplasmic marker GroEL confirmed that the fractionation was performed correctly. Since no putative signal peptide was predicted, it was concluded that the truncated RvCA5 was not the result of signal peptide cleavage.

To determine the cleavage site, purified RvCA5 was analyzed by MALDI-TOF mass spectrometry. The analysis yielded two major peaks (Figure 2d): one at approximately m/z 40,145, corresponding to the full-length RvCA5, and the other at m/z 37,374, corresponding to the truncated RvCA5. This mass difference indicated that the cleavage occurred specifically between Leucine 25 and Alanine 26. The protease responsible for this specific cleavage, however, remains to be elucidated.

2.4. Intrinsically Disordered Regions of RvCA5

For structural analysis, a predicted structure of the full-length RvCA5 was obtained using AlphaFold (Figure 3a). Consistent with the predicted classification as an α -CA, the structure of RvCA5 was highly similar to those of other α -CAs, such as bCA2, although the core antiparallel β -sheet consisted of 8 strands instead of the typical 10 strands [18]. Notably, the terminal regions exhibited unstructured extensions, which corresponded well to the N- and C-terminal extension sequences (Figure 1). Since the terminal extensions were likely intrinsically disordered regions (IDRs), the intrinsic disorder of RvCA5 was assessed using multiple IDR prediction tools via the CAID prediction portal. As expected, the N- and C-terminal regions were strongly predicted to be disordered, a finding that contrasts sharply with the structure of bCA2 (Figure 3b). The long and unstructured nature of the N-terminal region likely accounts for its susceptibility to proteolytic cleavage. Intriguingly, not only RvCA5 but also all other RvCA isoforms were predicted to possess extended terminal IDRs (Supplementary Figure S2), suggesting that this is a shared characteristic common to all RvCA isoforms.

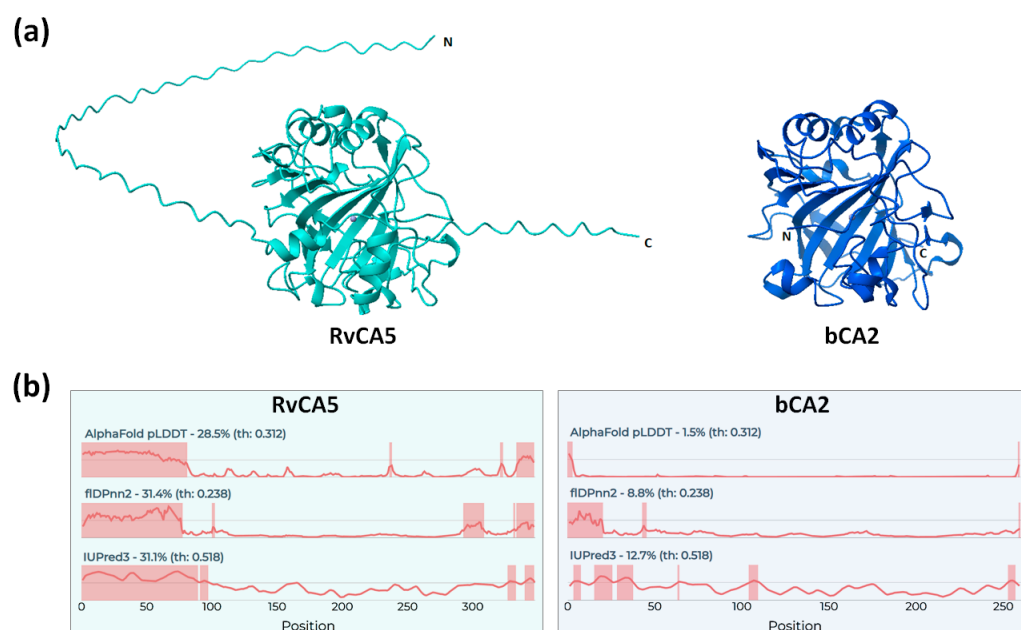


Figure 3. Intrinsically disordered regions (IDRs) of RvCA5. (a) Comparison of the RvCA5 structure with the bCA2 structure (PDB ID: 1V9E). (b) Sequence-based prediction of IDRs. IDRs are highlighted by shading.

2.5. Activity and Stability of RvCA5

The activity of RvCA5 was determined by the CO₂ hydration assay. The RvCA5 showed CO₂ hydration activity of 123.4 ± 22.0 WAU mg^{−1} (Figure 4a), which is only 1% of the activity of bCA2 ($12,700$ WAU mg^{−1}) [14]. When the RvCA5 was prepared in a buffer containing 300 mM NaCl, the activity decreased by an order of magnitude (to 13.0 ± 0.6 WAU mg^{−1}) compared to the activity of RvCA5 prepared in a buffer without NaCl supplementation (Figure 4a). This result suggests that RvCA5 is susceptible to salt-induced denaturation. The ICP-OES analysis showed that recombinant RvCA5 contains approximately 0.54 moles of zinc per mole of enzyme. Since Zn²⁺ coordination within the α -CA active site is strictly governed by the precise geometric arrangement of three conserved histidine residues [22], the observed 54% zinc occupancy confirms that the enzyme was properly folded. This indicates that the remarkably low activity is not a consequence of structural instability or a deficiency of the zinc cofactor, but is instead an intrinsic functional characteristic of RvCA5. Given its extremely low CO₂ hydration activity, RvCA5 is unlikely to function primarily in CO₂ metabolism, unlike typical CA enzymes.

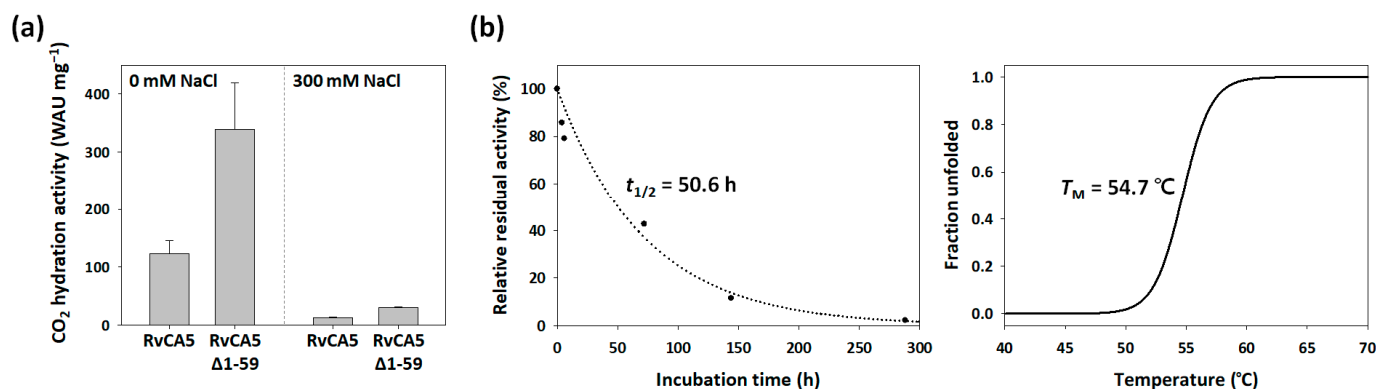


Figure 4. Activity and stability of RvCA5. (a) CO₂ hydration activity of RvCA5 and its derivative RvCA5Δ1-59. Enzymes were prepared in 20 mM sodium phosphate buffer (pH 7.5) with or without supplementation of 300 mM NaCl. Error bars represent standard deviations from at least three independent experiments. (b) Stability of RvCA5. Kinetic stability (left) was measured based on enzymatic CO₂ hydration activity after incubation at 40 °C. Data points (closed circles) were fitted to an exponential decay curve to determine the half-life ($t_{1/2}$). The thermodynamic stability (right) was estimated by temperature-dependent circular dichroism spectroscopy.

Moreover, the RvCA5 showed moderate stability with a half-life of 50.6 h at 40 °C and a melting temperature of 54.7 °C (Figure 4b), contradicting the initial anticipation that it would be an extremozyme. Recent studies have demonstrated that tardigrades in their active hydrated state are not inherently thermostable, and their exceptional thermal tolerance is primarily a feature of the desiccated state [23,24]. Thus, individual enzymes in tardigrades are likely susceptible to thermal stress, and the stability results for RvCA5 are consistent with the organism's actual physiological limits in its active state.

2.6. Expanding the Substrate Tunnel Improves the Activity of RvCA5

Next, the substrate tunnel of the enzyme was analyzed using Caver Web 2.0. α -CAs generally possess a deep, cone-shaped tunnel through which diffusion of substrates and products into and out of the buried active site occurs. The typical tunnel architecture of α -CA is also present in bCA2, manifesting as a deep (~10 Å) cavity extending close to the active site (Figure 5a) [25]. In contrast, a large area of the tunnel entrance on RvCA5 was blocked by a part of the N-terminal IDR, thereby reshaping it into three narrow and longer tunnels instead of a single wide one (Figure 5b). While this structural analysis is based on a

static snapshot, it is important to consider the highly dynamic nature of IDRs. Through rapid conformational fluctuations, the IDR can occupy a large excluded volume around the tunnel entrance [26], acting as a dynamic entropic barrier that hinders the diffusion of CO₂ molecules into the active site.

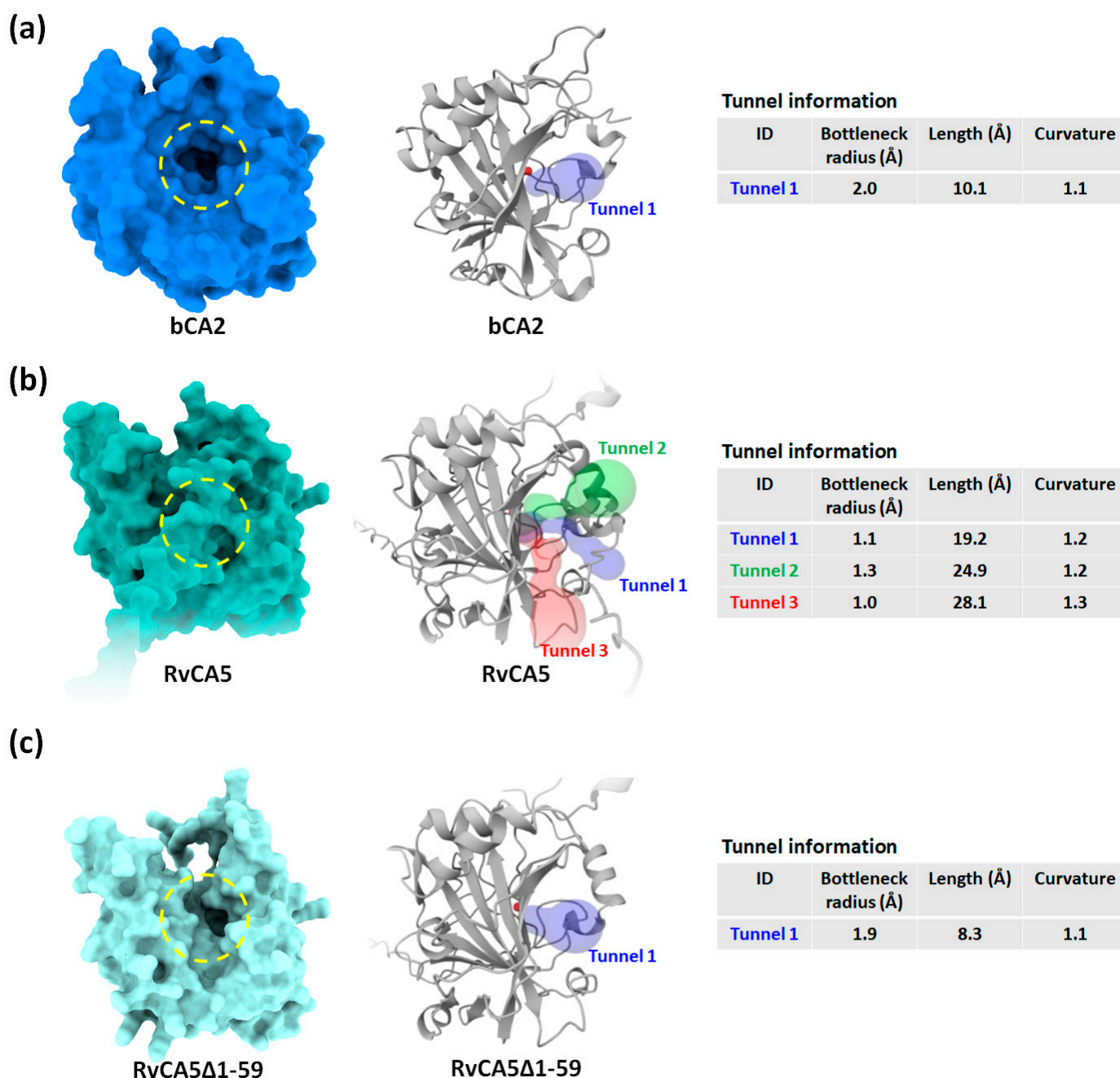


Figure 5. Analysis of substrate tunnel in (a) bCA2, (b) RvCA5, and (c) RvCA5Δ1-59, an N-terminal IDR-deleted derivative of RvCA5. The yellow dashed circle represents the cone-shaped cavity.

Based on the hypothesis that the extremely low activity of RvCA5 was (at least partially) due to inefficient mass transfer, an N-terminal IDR-deleted version of the enzyme (RvCA5Δ1-59) was constructed to test for improved activity. The removal of the blocking region successfully reverted the tunnel architecture to one closely resembling that of bCA2, featuring a much larger bottleneck radius (1.9 Å) and a shorter depth (8.3 Å) when compared to the original RvCA5 (Figure 5c). The CO₂ hydration activity of RvCA5Δ1-59 was more than two-fold higher compared to the original RvCA5, regardless of NaCl supple-

mentation (Figure 4a). Furthermore, the zinc content of RvCA5Δ1-59 was determined to be approximately 0.58 moles of zinc per mole of enzyme, a value closely comparable to the 0.54 mol/mol measured for the original RvCA5. These results collectively demonstrate that the significant increase in enzymatic activity upon the removal of the N-terminal IDR is primarily attributable to enhanced mass transfer, rather than any variation in zinc occupancy within the active site.

2.7. Unusual Surface Exposure of Cysteines

Many α -CAs, particularly of bacterial origin, have a conserved intramolecular disulfide bond (Figure 1) [12]. The other nine RvCAs, except for RvCA5, also possess the two corresponding cysteine residues required to form this conserved disulfide bond (Table 1). However, RvCA5 is distinct from the other RvCAs in that it lacks the two cysteine residues for the conserved disulfide bond and has a total of six cysteine residues, an exceptionally large number found in α -CAs. The six cysteines are spatially separated in the RvCA5 structure, preventing the formation of an intramolecular disulfide bond (Figure 6a).

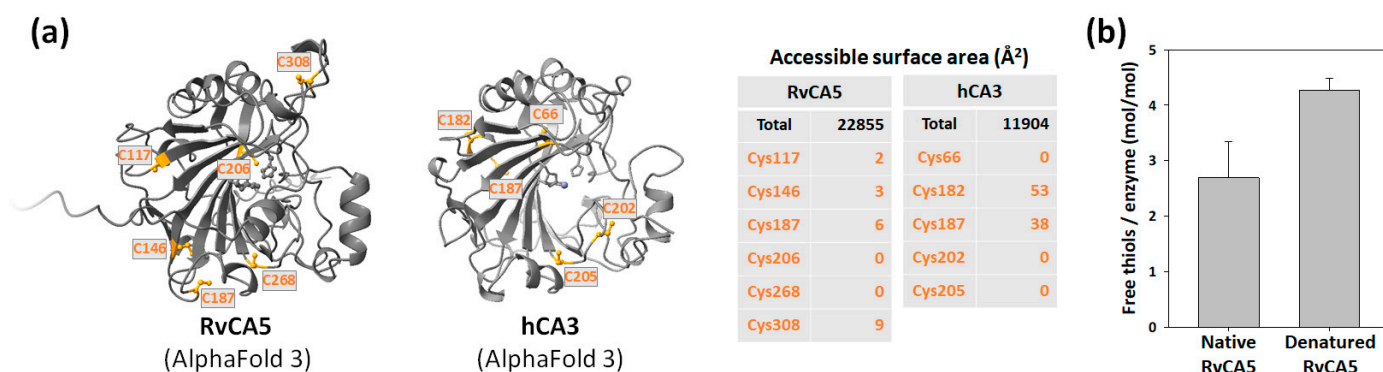


Figure 6. Assessment of surface-exposed cysteines. (a) Localization of surface-exposed cysteines on RvCA5 and hCA3. Residues are indicated in orange, and corresponding solvent-accessible surface area values are tabulated. (b) Quantitative analysis of free thiols in RvCA5 under native and denatured conditions using Ellman's assay.

The cytosolic human CA 3 (hCA3) has a very low catalytic activity (<1% of the fastest isozyme hCA2) [27]. Notably, hCA3 possesses five cysteine residues, among which two reactive cysteines are surface-exposed and susceptible to oxidative modifications such as S-glutathionylation (Figure 6a) [28,29]. hCA3 is believed to function as an antioxidant agent, utilizing its reactive cysteines to protect cells from oxidative damage. In the case of RvCA5, four of the six cysteine residues were predicted to be surface-exposed, albeit the corresponding solvent-accessible surface areas were calculated to be smaller than those of hCA3 (Figure 6a). Quantification of free thiols using Ellman's assay revealed that native RvCA5 possesses 2.7 free thiols per enzyme, confirming the presence of surface-exposed cysteine residues (Figure 6b). In its denatured state, the number of free thiols increased to 4.3 per enzyme. This value is lower than the theoretical total of 6 cysteines, likely due to partial oxidation of surface cysteines, resulting in the formation of intermolecular disulfide bonds. The observed difference in thiol content between the native and denatured states is presumably attributable to the two buried cysteines, Cys206 and Cys268. The presence of multiple surface-exposed cysteines raises the intriguing possibility that RvCA5 may function as a potential sensor that responds to changes in cellular redox status. Notably, the tardigrade homologs HeCA5A and PmCA also possess multiple surface-exposed cysteines (Supplementary Figure S1a,b). Furthermore, these three tardigrade CAs share two conserved, surface-exposed cysteine residues (Cys187 and Cys308 in RvCA5 numbering),

implying a potential functional conservation of redox-sensitive regulation across these tardigrade species.

2.8. Interaction of RvCA5 with DNA

Although RvCA5 is a slightly acidic protein with a pI value of 6.85, the analysis of surface electrostatic potential revealed a positively charged surface patch near the C-terminal IDR (Figure 7a). Since an NLS was identified in RvCA5, it was assumed that the positively charged surface patch might be involved in DNA binding. Accordingly, the binding of RvCA5 to DNA was tested in silico using the AlphaFold server with an arbitrary dsDNA sequence (5'-ATACCTAGGAACATTAGATACAGGGCCTTAGATGACTCTGATGCTAGCTAGC-3'). Surprisingly, the prediction showed that RvCA5 could bind to DNA via a positively charged surface patch (Figure 7b). Consistent results were observed across multiple random DNA sequences. In contrast, bCA2 lacks such a positively charged patch and was predicted not to bind to DNA (Figure 7a,b). TaCA, a bacterial periplasmic α -CA from *Thermovibrio ammonificans* that possesses positively charged surface patches, was also tested to exclude the possibility that the predicted binding of RvCA5 to the DNA was due to nonspecific electrostatic interactions. The result showed that TaCA did not interact with the DNA (Figure 7a,b), suggesting that the predicted DNA binding was specific to RvCA5. In RvCA5, the majority (5 out of 6) of the residues found within 4 Å of DNA were positively charged Lys and Arg residues that are crucial for electrostatic attraction to negatively charged DNA (Figure 7c). Notably, the C-terminal IDR extending near the DNA has residues that are potentially important for DNA binding: positively charged residues (R335, H336, K339, H341, K344) for electrostatic interactions with DNA backbone, and polar and nonpolar residues (Q334, E337, T340, Q342, V343, F345, T346, S347) for sequence recognition and binding specificity. Regarding the tardigrade homologs, PmCA was predicted to bind DNA, whereas HeCA5A was not (Supplementary Figure S1c,d). Given that no NLS was identified in the PmCA sequence, its predicted DNA-binding activity warrants further validation in future studies.

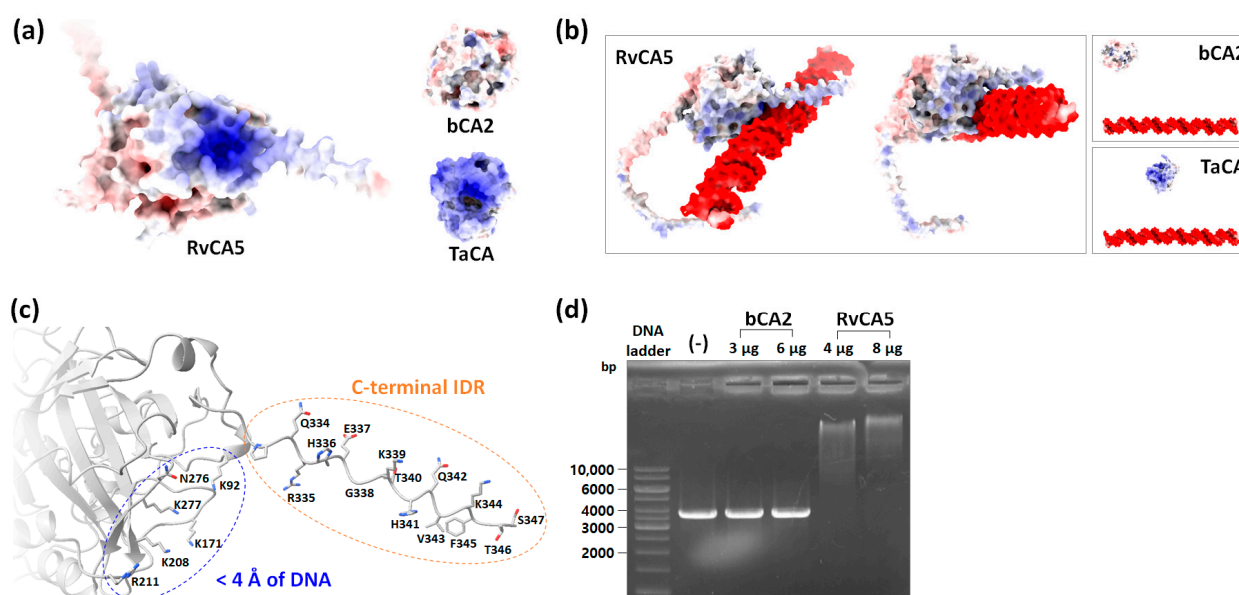


Figure 7. Interaction of RvCA5 and DNA. (a) Electrostatic surface potential map of CAs at pH 7, represented on a scale from -10 kT/e (red) to $+10$ kT/e (blue). (b) Prediction of DNA binding using AlphaFold. (c) Residues in RvCA5 that are potentially involved in DNA interaction. The six residues within 4 Å of DNA and the 15 residues of the C-terminal IDR are highlighted in blue and orange, respectively. (d) Experimental validation of DNA binding by EMSA. (-), negative control with DNA only.

To experimentally validate RvCA5's predicted DNA-binding activity, an electrophoretic mobility shift assay (EMSA) was performed. Incubation with purified RvCA5 induced a concentration-dependent migration shift in the linearized plasmid DNA (3741 bp). In contrast, bCA2 showed no interaction with the DNA substrate under the same conditions (Figure 7d). These results demonstrate a physical interaction between RvCA5 and DNA, although the specific physiological target sequences remain to be identified.

2.9. Putative Cellular Function of RvCA5

The structural and functional profile of RvCA5 identified in this study—namely, its low catalytic activity, multiple surface-exposed cysteines, terminal IDRs, and DNA-binding capability—suggests a cellular role distinct from typical enzymatic functions. RvCA5 may have evolved as a moonlighting protein by repurposing the ancestral CA scaffold for alternative functions [30]. The combination of low catalytic activity and surface cysteines closely parallels the characteristics of hCA3, a known cellular redox sensor [28,29]. Based on these structural features, a hypothesis is proposed that RvCA5 may function as a potential redox-responsive transcriptional regulator. In this model, RvCA5 is putatively translocated into the nucleus upon redox activation, where it may bind DNA via its positively charged surface patch. The C-terminal IDR might contribute to DNA-binding specificity through a disorder-to-order transition, while the N-terminal IDR could potentially serve as a flexible scaffold for cytoplasmic redox partners or recruit transcription factors to modulate gene expression. The abundant, constitutive expression of RvCA5 likely ensures an immediate functional response to environmental stresses. Given that terminal IDRs are common across the RvCA isoforms, these IDRs may also help stabilize proteins under extreme desiccation, as do other TDPs. Intriguingly, although the homologs HeCA5A and PmCA exhibit predictive profiles for NLS and DNA-binding ability that differ from those of RvCA5, they all share multiple surface-exposed cysteines. This conservation suggests that while the redox-sensing scaffold may be maintained across the tardigrade species, downstream mechanisms may have diverged or specialized across different tardigrade lineages. Although this putative mechanism remains to be experimentally confirmed, it provides a plausible framework for future in vivo validation of the physiological roles of RvCA5 in *R. varieornatus* and other tardigrade homologs.

3. Materials and Methods

3.1. Vector Construction

The *Escherichia coli* TOP10 strain was used for the construction and amplification of plasmid vectors, and the *E. coli* BL21(DE3) strain was used for the overexpression of recombinant proteins. *E. coli* cells were cultivated in a shaking incubator (Jeiotech, Daejeon, Republic of Korea) at 37 °C and 180 rpm in Luria–Bertani (LB) medium, added with 50 µg/mL ampicillin whenever necessary. For recombinant expression of RvCA5 (GenBank accession number: GAU95394), the previously constructed pET-RvCA5 vector was used, in which the RvCA5 gene is inserted into the pET-22b(+) vector via *Nde*I and *Xho*I restriction sites [17]. For the expression of the truncated version of RvCA5 (Δ 1–59), the corresponding DNA fragment was obtained from the pET-RvCA5 vector by PCR using forward primer 5'-CATATGCGTATGAACATGGCGC-3' and reverse primer 5'-CTCGAGGCTGGTAAATTTC-3' (*Nde*I and *Xho*I restriction sites underlined), and was subcloned into the pET-22b(+) vector, resulting in the pET-trncRvCA5 vector. For the expression of bovine CA 2 (bCA2), the bCA gene was chemically synthesized (Genscript, Piscataway, NJ, USA) and subcloned into pET-22b(+) using *Nde*I and *Xho*I restriction sites, yielding pET-bCA. All the genes had the His₆-tag sequence at their 3' terminus.

3.2. Expression and Purification of Recombinant Enzymes

The constructed vectors were each introduced into *E. coli* BL21(DE3), and the recombinant strains were cultured at 37 °C in LB medium supplemented with ampicillin. When the optical density at 600 nm (OD_{600}) of the culture reached 0.6 as measured via a spectrophotometer (UV-1800; Shimadzu, Kyoto, Japan), 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG; Duchefa Biochemie, Haarlem, The Netherlands) was added to the medium to induce recombinant protein expression driven by the T7lac promoter. Additionally, 0.1 mM $ZnSO_4$ (Junsei, Kyoto, Japan) was added to supplement Zn^{2+} , the cofactor required for CA activity. Following 12 h of post-induction cultivation at 25 °C, the cells were harvested by centrifugation at 4 °C and $4000 \times g$ for 10 min. The cells were resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, and 10 mM imidazole; pH 8.0) and subsequently disrupted using an ultrasonic homogenizer (Sonics and Materials, Newtown, CT, USA) while kept on ice. Then, the lysate was centrifuged at 4 °C and $10,000 \times g$ for 10 min, and the supernatant and the remaining debris were designated the soluble and insoluble fractions, respectively. For the affinity purification of recombinant enzymes, the soluble fraction was mixed with Ni^{2+} -nitrilotriacetic acid agarose beads (Qiagen, Germantown, MD, USA), and the His₆-tagged recombinant enzyme was purified according to the manufacturer's instructions by sequentially applying wash buffer (50 mM sodium phosphate, 300 mM NaCl, and 30 mM imidazole; pH 8.0) and elution buffer (50 mM sodium phosphate, 300 mM NaCl, and 250 mM imidazole; pH 8.0). Finally, the eluates were thoroughly dialyzed with a dilution factor of at least 10^{-11} against enzyme buffer (20 mM sodium phosphate; pH 7.5) at 4 °C using a Spectra/POR dialysis membrane with a 6–8 kDa MWCO (Spectrum Labs, Compton, CA, USA). If necessary, the enzyme buffer was supplemented with 300 mM NaCl.

3.3. SDS-PAGE and Western Blot

Protein samples were mixed with 4 $\times$ Laemmli sample buffer (Bio-Rad, Hercules, CA, USA) and heated for 10 min at 100 °C. They were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 15% PAGE gel and visualized by Coomassie blue R-250 (Bio-Rad, Hercules, CA, USA) staining. For periplasmic fractionation, the harvested cells were partially lysed and separated into the periplasmic fraction (Peri) and the spheroplast fraction (Sp) using the osmotic shock method as previously described [31]. For Western blot analysis, proteins separated by SDS-PAGE were transferred to a nitrocellulose membrane (Whatman, Piscataway, NJ, USA), and primary and secondary antibodies were applied sequentially. Mouse monoclonal anti-His₆ antibody (ABM, Vancouver, BC, Canada), rabbit polyclonal anti- β -lactamase antibody (Merck, Rahway, NJ, USA), and rabbit polyclonal anti-GroEL antibody (Enzo Life Sciences, Farmingdale, NY, USA) were used as the primary antibodies. Rabbit polyclonal anti-mouse IgG and goat anti-rabbit IgG antibodies (Bethyl Laboratories, Montgomery, TX, USA), both conjugated to alkaline phosphatase, were used as the secondary antibodies. Alkaline phosphatase activity was chromogenically detected using nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (Surmodics, Eden Prairie, MN, USA).

3.4. Mass Spectrometry

The purified RvCA5 enzyme solution was thoroughly dialyzed against deionized water before mass spectrometric analysis. The molecular mass of the purified RvCA5 was analyzed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Autoflex Speed; Bruker Daltonics, Billerica, MA, USA), operated in linear mode.

3.5. Protein Quantification

For protein quantification, the purified enzymes were denatured in denaturing buffer (6 M guanidine hydrochloride (GuHCl)/20 mM sodium phosphate buffer; pH 7.5), and the absorbance of the denatured protein was measured at 280 nm in a quartz crystal cuvette (Hellma Analytics, Müllheim, Germany). The protein concentration was determined from the measured absorbance using the molar extinction coefficient of RvCA5 at 280 nm, calculated by ProtParam (<http://web.expasy.org/protparam/>, accessed on 22 June 2025) [32].

3.6. CA Activity Assay

CA activity was measured using a colorimetric Wilbur-Anderson CO₂ hydration assay performed at 0 °C within a UV-1800 spectrophotometer equipped with a temperature-controlled cell holder, as previously described [14]. Briefly, 5–50 µL of the enzyme sample at 20 µM was added to a disposable cuvette (Sarstedt, Lower Saxony, Germany) containing 600 µL of assay buffer (20 mM Tris, 100 µM phenol red; pH 8.3). The CO₂ hydration reaction was initiated by adding 400 µL of CO₂-saturated deionized water, which was prepared in ice-cold water, to the cuvette. The absorbance change was monitored at 570 nm, and the time (*t*) required for the pH drop from 7.5 to 6.5 was obtained. The time (*t*₀) for the uncatalyzed reaction was also measured by adding the same volume of the corresponding blank buffer. The Wilbur-Anderson unit (WAU) was calculated as (*t*₀ − *t*)/*t*.

3.7. Zinc Content Analysis

The zinc content of purified enzyme was quantified by inductively coupled plasma optical emission spectroscopy (ICP-OES) on an Optima 8300 DV instrument (PerkinElmer, Waltham, MA, USA).

3.8. Thermostability Test

For the kinetic stability test, the samples were heated at 40 °C in a water bath (Jeiotech, Daejeon, Republic of Korea) for up to 12 days. They were then stored at 4 °C until enzyme activity was measured using the CO₂ hydration assay. The relative residual activity was calculated based on the activity of the unheated sample as follows:

$$\text{Relative residual activity(\%)} = \left(\frac{\text{WAU of heated sample}}{\text{WAU of unheated sample}} \right) \times 100$$

For the thermodynamic stability test, a circular dichroism (CD) spectrum was obtained using a CD spectropolarimeter (Jasco, Tokyo, Japan). The temperature of the RvCA5 solution inside a quartz cuvette was increased from 25 °C to 70 °C at a rate of 2 °C/min by heating. The change in the ellipticity was recorded at 220 nm, and the obtained data were fitted to a Boltzmann sigmoidal equation using SigmaPlot 10.0 (Systat Software, San Jose, CA, USA).

3.9. Free Thiol Quantification

Free thiol content was measured according to Ellman's assay [33]. First, 250 µL of RvCA5 solution was added to a reaction solution containing 700 µL of denaturing buffer. Subsequently, 50 µL of 2 mM 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB; Sigma-Aldrich, St. Louis, MO, USA), dissolved in 20 mM sodium phosphate buffer (pH 7.5), was added to the reaction solution. After the reaction was allowed to proceed for 5 min, the mixture was transferred to a quartz cuvette (Hellma Analytics), and the absorbance was measured at 412 nm. For blank correction, 20 mM sodium phosphate buffer (pH 7.5) was added instead of the enzyme sample. The measured absorbance was used to calculate the free thiol concentration using the extinction coefficient (13,600 M^{−1} cm^{−1}) of 2-nitro-5-thiobenzoic

acid. For measurement under native conditions, the same volume of 20 mM sodium phosphate buffer (pH 7.5) was used instead of the denaturing buffer.

3.10. Electrophoretic Mobility Shift Assay (EMSA)

A pGEM-T easy vector (Promega, Madison, WI, USA) harboring the superfolder GFP gene flanked by a HindIII site was linearized using the HindIII restriction enzyme (ThermoFisher Scientific, Waltham, MA, USA) to generate a 3741 bp dsDNA. The linearized DNA was purified via agarose gel electrophoresis using the FavorPrepTM Gel/PCR purification kit (FAVORGEN, Ping Tung, Taiwan) and eluted in deionized water. For the binding assay, 10 μ L of the DNA solution (approximately 200 ng) was mixed with an equal volume of the enzyme solution (10 μ M or 20 μ M) prepared in 20 mM sodium phosphate buffer (pH 7.5), and the mixture was incubated for 1 h at room temperature. The samples were then mixed with ExcelDyeTM 6 $\times$ DNA loading dye (SMOBIO, Hsinchu City, Taiwan) and resolved on a 1% agarose gel with TAE buffer (40 mM Tris, 1.14 mL glacial acetic acid, 1 mM EDTA) and subsequently stained with StaySafeTM Nucleic Acid Stain (RBC Bioscience, New Taipei City, Taiwan). DNA bands were visualized using the GelDoc Go imaging system (Bio-Rad, Hercules, CA, USA).

3.11. In Silico Analyses

Multiple sequence alignment was performed using ClustalX2.0 [34] and visualized using Jalview 2.11.5.1 [35]. Signal peptide prediction was performed using SignalP 6.0 [36]. Mitochondrial targeting sequence was predicted using TargetP 2.0 [37], DeepMito [38], and Euk-mPLoc 2.0 [39]. Nuclear localization signal was identified using DeepLoc 2.1 [40]. The translation initiation region was predicted using the UTR Designer [41]. Protein structure was predicted using the AlphaFold 3 server [42] and visualized with UCSF ChimeraX 1.9 [43]. Intrinsically disordered regions (IDRs) and binding regions inside IDRs were predicted using the CAID prediction portal [44]. Analysis of tunnels in protein structures was performed using the CAVER Web server [45] with default values for CAVER parameters, except for the following parameters: shell depth (20 Å), shell radius (10 Å), and clustering threshold (5 Å). The solvent-accessible surface area was assessed using the XSSP server [46]. Electrostatic potential on protein surface was calculated by Adaptive Poisson–Boltzmann Solver (APBS) [47].

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/molecules31030538/s1>. Figure S1: Comparison of RvCA5 with homologous tardigrade CAs. HeCA5A, carbonic anhydrase from *Hypsibius exemplaris*; PmCA, carbonic anhydrase from *Paramacrobiotus metropolitanus*. (a) Multiple sequence alignments. Cysteine residues are boxed in yellow, and the conserved, surface-exposed cysteines (▼) are marked. (b) Assessment of surface-exposed cysteines. Cysteine residues are indicated in orange, and corresponding solvent-accessible surface area values are tabulated. The conserved, surface-exposed cysteines are marked in red. (c) Electrostatic surface potential map of CAs at pH 7, represented on a scale from −10 kT/e (red) to +10 kT/e (blue). (d) Prediction of DNA binding using AlphaFold. Figure S2. Structure of RvCA isoforms predicted by AlphaFold.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Ministry of Science & ICT (RS-2024-00355356) and by the Glocal University 30 Project Fund of Gyeongsang National University in 2025.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Hashimoto, T.; Horikawa, D.D.; Saito, Y.; Kuwahara, H.; Kozuka-Hata, H.; Shin, I.T.; Minakuchi, Y.; Ohishi, K.; Motoyama, A.; Aizu, T.; et al. Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade-unique protein. *Nat. Commun.* **2016**, *7*, 12808. [\[CrossRef\]](#)
2. Richaud, M.; Shrivastava, R.; Xirodimas, D.; Galas, S.; Bailly, A. Tardigrade Dsup extends *C. elegans* life span by impeding mitochondrial respiration and promoting oxidative stress resistance. *Sci. Adv.* **2025**, *11*, eadx9669. [\[CrossRef\]](#)
3. Aguilar, R.R.; Khan, L.F.; Cummins, C.K.; Arslanovic, N.; Grauer, T.; Birmingham, K.; Kasliwal, K.; Posnikoff, S.D.L.; Chakraborty, U.; Hickman, A.R.; et al. Multivalent binding of the tardigrade Dsup protein to chromatin promotes yeast survival and longevity upon exposure to oxidative damage. *Nat. Commun.* **2025**, *16*, 8617. [\[CrossRef\]](#)
4. Kirke, J.; Jin, X.L.; Zhang, X.H. Expression of a tardigrade Dsup gene enhances genome protection in plants. *Mol. Biotechnol.* **2020**, *62*, 563–571. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Zarubin, M.; Azorskaya, T.; Kuldoshina, O.; Alekseev, S.; Mitrofanov, S.; Kravchenko, E. The tardigrade Dsup protein enhances radioresistance in *Drosophila melanogaster* and acts as an unspecific repressor of transcription. *iScience* **2023**, *26*, 106998. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Yagi-Utsumi, M.; Aoki, K.; Watanabe, H.; Song, C.; Nishimura, S.; Satoh, T.; Yanaka, S.; Ganser, C.; Tanaka, S.; Schnapka, V.; et al. Desiccation-induced fibrous condensation of CAHS protein from an anhydrobiotic tardigrade. *Sci. Rep.* **2021**, *11*, 21328. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Miyazawa, K.; Itoh, S.G.; Watanabe, H.; Uchihashi, T.; Yanaka, S.; Yagi-Utsumi, M.; Kato, K.; Arakawa, K.; Okumura, H. Tardigrade secretory-abundant heat-soluble protein has a flexible β -barrel structure in solution and keeps this structure in dehydration. *J. Phys. Chem. B* **2021**, *125*, 9145–9154. [\[CrossRef\]](#)
8. Kato, S.; Deguchi, K.; Obana, M.; Fujio, Y.; Fukuda, Y.; Inoue, T. Metabolite phosphatase from anhydrobiotic tardigrades. *FEBS J.* **2024**, *291*, 5195–5213. [\[CrossRef\]](#)
9. Law, Y.S.; Muthaliff, N.A.; Wei, Y.F.; Lin, F.; Zhao, H.M.; Ang, E.L. Biochemical investigation and engineering of a tardigrade X family DNA polymerase for template-independent DNA synthesis. *ACS Catal.* **2024**, *14*, 12318–12330. [\[CrossRef\]](#)
10. Esquirol, L.; Newman, J.; Nebl, T.; Scott, C.; Vickers, C.; Sainsbury, F.; Peat, T.S. Characterization of novel mevalonate kinases from the tardigrade *Ramazzottius varieornatus* and the psychrophilic archaeon *Methanococcoides burtonii*. *Acta Crystallogr. D Struct. Biol.* **2024**, *80*, 203–215. [\[CrossRef\]](#)
11. Sim, K.S.; Inoue, T. Structure of a superoxide dismutase from a tardigrade: *Ramazzottius varieornatus* strain YOKOZUNA-1. *Acta Crystallogr. F Struct. Biol. Commun.* **2023**, *79*, 169–179. [\[CrossRef\]](#)
12. Smith, K.S.; Ferry, J.G. Prokaryotic carbonic anhydrases. *FEMS Microbiol. Rev.* **2000**, *24*, 335–366. [\[CrossRef\]](#)
13. Luca, V.D.; Vullo, D.; Scozzafava, A.; Carginale, V.; Rossi, M.; Supuran, C.T.; Capasso, C. An α -carbonic anhydrase from the thermophilic bacterium *Sulphurihydrogenibium azorense* is the fastest enzyme known for the CO₂ hydration reaction. *Bioorg. Med. Chem.* **2013**, *21*, 1465–1469. [\[CrossRef\]](#)
14. Kim, J.H.; Jo, B.H. A colorimetric CO₂ hydration assay for facile, accurate, and precise determination of carbonic anhydrase activity. *Catalysts* **2022**, *12*, 1391. [\[CrossRef\]](#)
15. Talekar, S.; Jo, B.H.; Dordick, J.S.; Kim, J. Carbonic anhydrase for CO₂ capture, conversion and utilization. *Curr. Opin. Biotechnol.* **2022**, *74*, 230–240. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Lim, G.T.; Jo, B.H. Cation-assisted stabilization of carbonic anhydrase one-step in situ loaded in diatom-inspired silica nanospheres for potential applications in CO₂ capture and utilization. *Sep. Purif. Technol.* **2024**, *333*, 125926. [\[CrossRef\]](#)
17. Choi, Y.J.; Jo, B.H. Recombinant expression and purification of α -carbonic anhydrases from a water bear *Ramazzottius varieornatus* in *Escherichia coli*. *Korean J. Microbiol.* **2023**, *59*, 204–213.
18. Aggarwal, M.; Boone, C.D.; Kondeti, B.; McKenna, R. Structural annotation of human carbonic anhydrases. *J. Enzyme Inhib. Med. Chem.* **2013**, *28*, 267–277. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Yoshida, Y.; Koutsovoulos, G.; Laetsch, D.R.; Stevens, L.; Kumar, S.; Horikawa, D.D.; Ishino, K.; Komine, S.; Kunieda, T.; Tomita, M.; et al. Comparative genomics of the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus*. *PLoS Biol.* **2017**, *15*, e2002266. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Aspatwar, A.; Supuran, C.T.; Waheed, A.; Sly, W.S.; Parkkila, S. Mitochondrial carbonic anhydrase VA and VB: Properties and roles in health and disease. *J. Physiol.* **2023**, *601*, 257–274. [\[CrossRef\]](#)
21. Owji, H.; Nezafat, N.; Negahdaripour, M.; Hajiebrahimi, A.; Ghasemi, Y. A comprehensive review of signal peptides: Structure, roles, and applications. *Eur. J. Cell Biol.* **2018**, *97*, 422–441. [\[CrossRef\]](#)
22. Christianson, D.W.; Fierke, C.A. Carbonic anhydrase: Evolution of the zinc binding site by nature and by design. *Acc. Chem. Res.* **1996**, *29*, 331–339. [\[CrossRef\]](#)
23. Neves, R.C.; Hvidepil, L.K.B.; Sørensen-Hygum, T.L.; Stuart, R.M.; Møbjerg, N. Thermotolerance experiments on active and desiccated states of *Ramazzottius varieornatus* emphasize that tardigrades are sensitive to high temperatures. *Sci. Rep.* **2020**, *10*, 94. [\[CrossRef\]](#)

24. Loeffelholz, J.; Meese, E.; Giovannini, I.; Ullibbarri, K.; Momeni, S.; Merfeld, N.; Wessel, J.; Guidetti, R.; Rebecchi, L.; Boothby, T.C. An evaluation of thermal tolerance in six tardigrade species in an active and dry state. *Biol. Open* **2024**, *13*, bio060485. [\[CrossRef\]](#)
25. Saito, R.; Sato, T.; Ikai, A.; Tanaka, N. Structure of bovine carbonic anhydrase II at 1.95 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 792–795. [\[CrossRef\]](#)
26. Hoh, J.H. Functional protein domains from the thermally driven motion of polypeptide chains: A proposal. *Proteins* **1998**, *32*, 223–228. [\[CrossRef\]](#)
27. Duda, D.M.; Tu, C.; Fisher, S.Z.; An, H.; Yoshioka, C.; Govindasamy, L.; Laipis, P.J.; Agbandje-McKenna, M.; Silverman, D.N.; McKenna, R. Human carbonic anhydrase III: Structural and kinetic study of catalysis and proton transfer. *Biochemistry* **2005**, *44*, 10046–10053. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Di Fiore, A.; Monti, D.M.; Scaloni, A.; De Simone, G.; Monti, S.M. Protective role of carbonic anhydrases III and VII in cellular defense mechanisms upon redox unbalance. *Oxid. Med. Cell Longev.* **2018**, *2018*, 2018306. [\[CrossRef\]](#)
29. Yu, Y.; Poulsen, S.A.; Di Trapani, G.; Tonissen, K.F. Exploring the redox and pH dimension of carbonic anhydrases in cancer: A focus on carbonic anhydrase 3. *Antioxid. Redox Signal* **2024**, *41*, 957–975. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Huberts, D.H.E.W.; van der Klei, I.J. Moonlighting proteins: An intriguing mode of multitasking. *Biochim. Biophys. Acta* **2010**, *1803*, 520–525. [\[CrossRef\]](#)
31. Jo, B.H.; Kim, I.G.; Seo, J.H.; Kang, D.G.; Cha, H.J. Engineered *Escherichia coli* with periplasmic carbonic anhydrase as a biocatalyst for CO₂ sequestration. *Appl. Environ. Microbiol.* **2013**, *79*, 6697–6705. [\[CrossRef\]](#)
32. Wilkins, M.R.; Gasteiger, E.; Bairoch, A.; Sanchez, J.C.; Williams, K.L.; Appel, R.D.; Hochstrasser, D.F. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **1999**, *112*, 531–552.
33. Simpson, R.J. Estimation of free thiols and disulfide bonds using Ellman’s reagent. *CSH Protoc.* **2008**, *2008*, pdb.prot4699. [\[CrossRef\]](#)
34. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [\[CrossRef\]](#)
35. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gislason, M.H.; Pihl, S.I.; Tsigirgos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **2022**, *40*, 1023–1025. [\[CrossRef\]](#)
37. Almagro Armenteros, J.J.; Salvatore, M.; Emanuelsson, O.; Winther, O.; von Heijne, G.; Elofsson, A.; Nielsen, H. Detecting sequence signals in targeting peptides using deep learning. *Life Sci. Alliance* **2019**, *2*, e201900429. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Savojardo, C.; Bruciaferri, N.; Tartari, G.; Martelli, P.L.; Casadio, R. DeepMito: Accurate prediction of protein sub-mitochondrial localization using convolutional neural networks. *Bioinformatics* **2020**, *36*, 56–64. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Chou, K.C.; Shen, H.B. A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0. *PLoS ONE* **2010**, *5*, e9931. [\[CrossRef\]](#)
40. Odum, M.T.; Teufel, F.; Thumhuri, V.; Almagro Armenteros, J.J.; Johansen, A.R.; Winther, O.; Nielsen, H. DeepLoc 2.1: Multi-label membrane protein type prediction using protein language models. *Nucleic Acids Res.* **2024**, *52*, W215–W220. [\[CrossRef\]](#)
41. Seo, S.W.; Yang, J.S.; Kim, I.; Yang, J.; Min, B.E.; Kim, S.; Jung, G.Y. Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metab. Eng.* **2013**, *15*, 67–74. [\[CrossRef\]](#)
42. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A.J.; Bambrick, J.; et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630*, 493–500. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **2023**, *32*, e4792. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Del Conte, A.; Bouhraoua, A.; Mehdiabadi, M.; Clementel, D.; Monzon, A.M.; CAID predictors; Tosatto, S.C.E.; Piovesan, D. CAID prediction portal: A comprehensive service for predicting intrinsic disorder and binding regions in proteins. *Nucleic Acids Res.* **2023**, *51*, W62–W69. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Marques, S.M.; Borko, S.; Vavra, O.; Dvorsky, J.; Kohout, P.; Kabourek, P.; Hejtmanek, L.; Damborsky, J.; Bednar, D. Caver Web 2.0: Analysis of tunnels and ligand transport in dynamic ensembles of proteins. *Nucleic Acids Res.* **2025**, *53*, W132–W142. [\[CrossRef\]](#)
46. Touw, W.G.; Baakman, C.; Black, J.; te Beek, T.A.; Krieger, E.; Joosten, R.P.; Vriend, G. A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* **2015**, *43*, D364–D368. [\[CrossRef\]](#)
47. Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L.E.; Brookes, D.H.; Wilson, L.; Chen, J.; Liles, K.; et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **2018**, *27*, 112–128. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.