Structural and Dynamic Characterization of *Candida boidinii* Formate Dehydrogenase by High-Resolution X-ray Crystallography

Mehmet Gul [1,†], Busra Yuksel [1,2,†], Huri Bulut [3] and Hasan DeMirci [1,4,5,*]

[1] Department of Molecular Biology and Genetics, Koc University, 34450, Istanbul, Türkiye; mgul21@ku.edu.tr (M.G.); byuksel20@ku.edu.tr (B.Y.); hdemirci@ku.edu.tr (H.D.)
[2] Max Planck Institute for Biophysics, 60438, Frankfurt am Main, Germany; buesra.yueksel@biophys.mpg.de (B.Y.)
[3] Department of Medical Biochemistry, Faculty of Medicine, Istinye University, 34010, Istanbul, Türkiye; huri.bulut@istinye.edu.tr (H.B.)
[4] Koc University Isbank Center for Infectious Diseases (KUISCID), Koc University, 34010, Istanbul, Türkiye
[5] Stanford PULSE Institute, SLAC National Laboratory, Menlo Park, CA 94025, USA
* Correspondence: hdemirci@ku.edu.tr; hdemirci@stanford.edu
† Equal contribution

Abstract: *Candida boidinii* NAD^+-dependent formate dehydrogenase (*Cb*FDH) has gained significant attention for its potential applications in the production of biofuels and various industrial chemicals from inorganic carbon dioxide. In this study, we present an atomic X-ray crystal structure of the apo *Cb*FDH to 1.4 Å resolution determined at cryogenic temperature at the Turkish Light Source, *"Turkish DeLight"*. This structure offers a comprehensive view of the apo enzyme's dynamics, filling the gaps in our understanding from previous studies. Also, comparison of our high-resolution apo and previously available holo enzyme structures reveals major conformational changes of this dynamic enzyme in the absence and presence of the coenzyme and substrate/inhibitor complexes. Collectively all these information may provide invaluable insights into future protein engineering efforts that could enhance enzymatic activity and stability, potentially increasing its efficiency and effectiveness of *Cb*FDH in industrial processes.

Keywords: formate dehydrogenase; *Candida boidinii*; protein engineering; X-ray crystallography; structural biology; structural dynamics; Turkish Light Source; *Turkish DeLight*.

1. Introduction

NAD^+-dependent formate dehydrogenase (FDH) is 364 amino acid containing homodimeric 82 KDa enzyme complex that catalyzes the reversible conversion between formate and carbon dioxide [1]. It uses NAD^+/NADH coenzymes for electron transfer to catalyze the oxidation of formate to carbon dioxide and vice versa [2]. Many prokaryotic and eukaryotic organisms express NAD^+-dependent FDH enzymes with similar structural and kinetic properties [3].

However, many studies have focused on the *Candida boidinii* NAD^+-dependent FDH (*Cb*FDH) due to its remarkable stability and activity [4].

Formate dehydrogenase was first discovered in the 1950s and recently gained great attention as a research topic due to its potential for formate molecule production from inorganic carbon dioxide. Formate is a valuable single-carbon containing precursor molecule used in the synthesis of various invaluable industrial chemicals, such as formaldehyde which is used in the production of resins, plastics, and disinfectants [5-7]. Other chemicals that can be generated from formate include formic acid and oxalic acid, both of which have a range of important industrial applications [8]. Moreover, the formate molecule has the potential as an alternative biofuel since it can be converted into hydrogen gas by electrolysis [9]. Overall, formate's versatility in industrial applications has made the FDH enzyme a central research topic.

High-resolution structural studies are crucial for understanding the function, underpinnings of the catalytic mechanism, and other potential applications of dehydrogenase enzymes like *Cb*FDH. Previous studies have contributed to our knowledge of the *Cb*FDH enzyme's structural and functional features. However, there are still existing gaps in our understanding of FDH function and potential modifications for further enhanced activity and thermal stability of this enzyme. Here, we present a high-resolution crystal structure of the apo *Cb*FDH determined at cryogenic temperature. Our findings, together with the currently available structural and enzymatic knowledge, will pave the way for future research on the potential use of *Cb*FDH in industrial applications.

2. Materials and Methods

2.1. Transformation, expression, purification, and crystallization

Full-length wild-type *Cb*FDH gene was cloned into pET-23a (+) bacterial expression vector with a C-terminal hexahistidine-tag as previously described by Bulut *et al*. [4]. The plasmid was transformed into the competent *E. coli* BL21 Rosetta-2 strain. Transformed bacterial cells were grown in 6 L LB medium containing 100 µg/ml ampicillin and 35 µl/ml chloramphenicol at 37 ℃. Once OD_{600} reached to 0.7-0.8 range, the protein expression was induced by addition of β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM for 3-4 hours at 37 ℃.

Cells were then harvested using Beckman Allegra 15R desktop centrifuge at 4 ℃ at 3500 rpm for 30 minutes. Cell pellets were stored at -80 ℃ until purification.

The frozen cell pellets were dissolved in ice cold lysis buffer that contains 500 mM NaCl, 50 mM Tris-HCl pH 7.5, 5% (*v/v*) Glycerol, 0.1% (*v/v*) Triton X-100, and sonicated on ice with Branson W250 sonifier (Brookfield, CT, USA) until the solution was completely cleared. The cell lysate was centrifuged at 4 ℃ at 35000 rpm for 1 hour by Beckman Optima™ L-80XP ultracentrifuge equipped with a Ti-45 rotor (Beckman, USA). The cell pellet was discarded, and the supernatant containing the soluble protein was filtered through 0.2 µm membrane and loaded onto a Ni-NTA column that was previously equilibrated with 3 column volume of loading buffer containing 200 mM NaCl, 20 mM Tris-HCl pH 7.5, 20 mM Imidazole. Unbound and non-specifically bound proteins were washed by running 5 column volumes of loading buffer through the column. The hexahistidine-tagged *Cb*FDH proteins were eluted with the elution buffer containing 200 mM NaCl, 20 mM Tris-HCl pH 7.5, and 250 mM Imidazole. The eluted protein was dialyzed in a dialysis membrane (3 kDa MWCO) against a buffer containing 200 mM NaCl and 20 mM Tris-HCl pH 7.5 overnight at 4 ℃ to remove excess imidazole. The concentrated ultra-pure protein was run on an 15% SDS-PAGE for verification.

The protein crystallization screening experiments were performed using the sitting-drop microbatch under oil method. The purified *Cb*FDH protein was mixed with ~3500 commercially available sparse matrix crystallization screening conditions in a 1:1 volumetric ratio in 72-Terasaki plates (Greiner Bio-One, Kremsmünster, Austria). The mixtures were covered with 16.6 µl 100% paraffin oil (Tekkim Kimya, Istanbul, Türkiye). The crystallization plates were incubated at 4 ℃ and frequently checked under a light microscope. The best *Cb*FDH crystals were grown within six weeks in Wizard III condition #47 (Hampton Research, USA). This condition consists of 30% (*w/v*) PEG 5000 MME, 100 mM MES/Sodium hydroxide pH 6.5, and 200 mM Ammonium sulfate.

2.2. Crystal harvesting and delivery

The *Cb*FDH crystals were harvested by using MiTeGen MicroLoops mounted to a magnetic wand [10] while simultaneously monitoring under a light microscope, as described by Atalay *et*

al. [11]. The harvested crystals were flash-frozen by plunging them in liquid nitrogen and placed in a cryo-cooled robotic sample puck (Cat#M-CP-111-021, MiTeGen, USA). The puck was mounted to a puck transfer and mounting tool and placed into the sample dewar which is auto refilled with liquid nitrogen at 100 °K, of the *Turkish DeLight* XRD device.

2.3. Data collection and processing

Diffraction data from the *Cb*FDH crystals were collected using Rigaku's XtaLAB Synergy Flow XRD at University of Health Sciences (Istanbul, Türkiye) controlled with the *CrysAlis Pro* software version 1.171.42.35a [12]. The crystals were constantly cooled by the Oxford Cryosystem's Cryostream 800 Plus set to 100 °K. The PhotonJet-R X-ray generator was operated at 40 kV, 30 mA, 1200.0 W with a beam intensity of 10%. The data was collected with 1.54 Å wavelength and the detector distance was 47.00 mm. The scan width was set to 0.25 degrees while exposure time was 10.0 minutes. Data collection was performed for 15 hours 43 minutes 19 seconds. Data reduction was performed using the *CrysAlis Pro* version 1.171.42.35a. The data reduction result was obtained as an *.mtz file.

2.4. Structure determination and refinement

The cryogenic *Cb*FDH structure was determined at 1.4 Å with the space group P1 using the automated molecular replacement program *PHASER* version 2.8.3 [13] within the *PHENIX* suite version 1.20.1 [14]. The previously released FDH structure with PDB ID: 5DNA was used as an initial search model for molecular replacement [15]. Initially rigid-body and simulated-annealing refinements were performed by *PHENIX*. Then, individual coordinates and translation/libration/screw (TLS) parameters were refined along with simulated annealing map refinement. The final model of the *Cb*FDH structure was checked by using *COOT* version 0.9.8.1 [16] after each refinement cycle. Water molecules were added into appropriate electron density clouds while those located outside the density were manually removed. The structure figures were generated by using *PyMOL* [17] version 2.5.2 and *COOT* version 0.9.8.1. Morph analysis was performed and the movies were generated using *PyMOL*.

Table 1. Data collection and refinement statistics.

Dataset	***Cb*FDH**
PDB ID	8HTY
Data collection	
Beamline	*Turkish DeLight*
Space group	P 1
Cell dimensions	
a, b, c (Å)	54.132, 68.959, 109.79
α, β, γ (°)	78.139, 89.482, 81.115
Resolution (Å)	24.75-1.39 (1.51-1.39)
CC (½)	0.999 (0.340)
*CC**	1.000 (0.708)
I / σI	11.74 (0.73)
Completeness (%)	98.82 (86.0)
Redundancy	151.46 (148.36)
Refinement	
Resolution (Å)	21.83-1.40 (1.43-1.40)
No. reflections	301717 (29123)
R_{work} / Rf_{ree}	0.1674/0.1904 (0.3802/0.3728)
No. atoms	
Protein	22225
Ligand / Ion / Water	2249
B-factors	
Protein	22.94
Ligand / Ion / Water	35.07
Coordinate errors	0.18
R.m.s deviations	
Bond lengths (Å)	0.015
Bond angles (°)	1.422
Ramachandran plot	
Favored (%)	97.38
Allowed (%)	2.62
Disallowed (%)	-

[1] The highest resolution shell is shown in parentheses.

3. Results

3.1. Apo CbFDH structure is determined at 1.4 Å resolution at the Turkish Light Source

We determined the crystal structure of wild-type apo *Cb*FDH at 1.4 Å resolution (PDB ID: 8HTY) with overal 98.82% completeness at cryogenic temperature at the Turkish Light Source "*Turkish DeLight*" (Istanbul, Türkiye) [18]. The determined structure consists of two homodimers of *Cb*FDH (Figure 1). The Ramachandran plot indicates that 98.01% of the residues were in the favored regions while 1.99% were in the allowed regions with no outliers. The data collection and refinement statistics are given in Table 1.

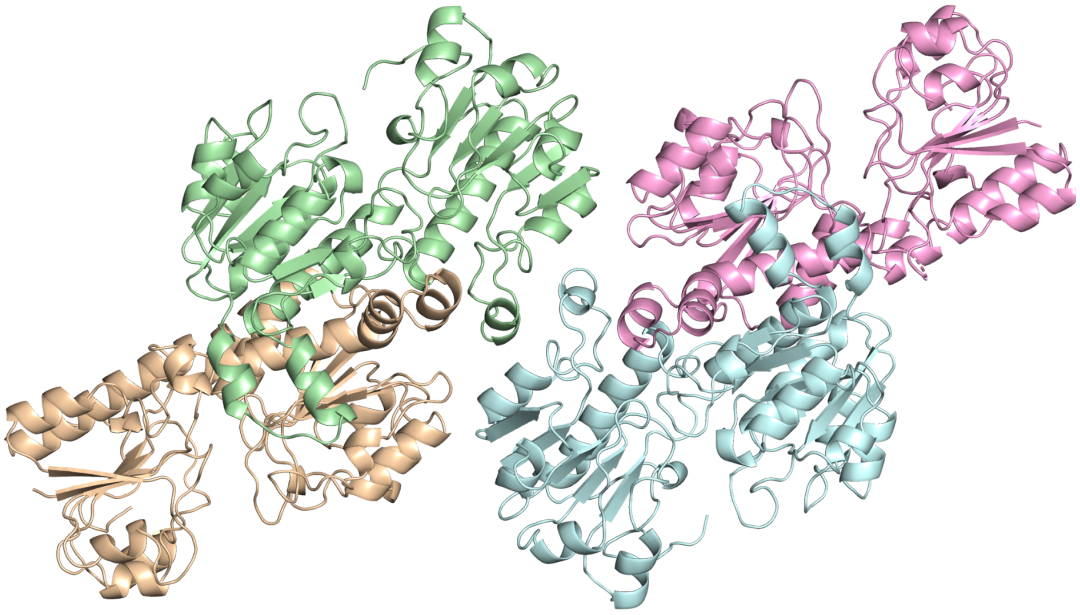

Figure 1. 1.4 Å apo wild-type *Cb*FDH structure as two homodimers. Four monomer molecules are shown in different colors.

The resulting final electron density map was superior-quality and reveals the structural features such as coordinated water molecules, and side chains of the *Cb*FDH molecule in great details. The electron density remains intact and visible even at 3 sigma levels (Figure 2).

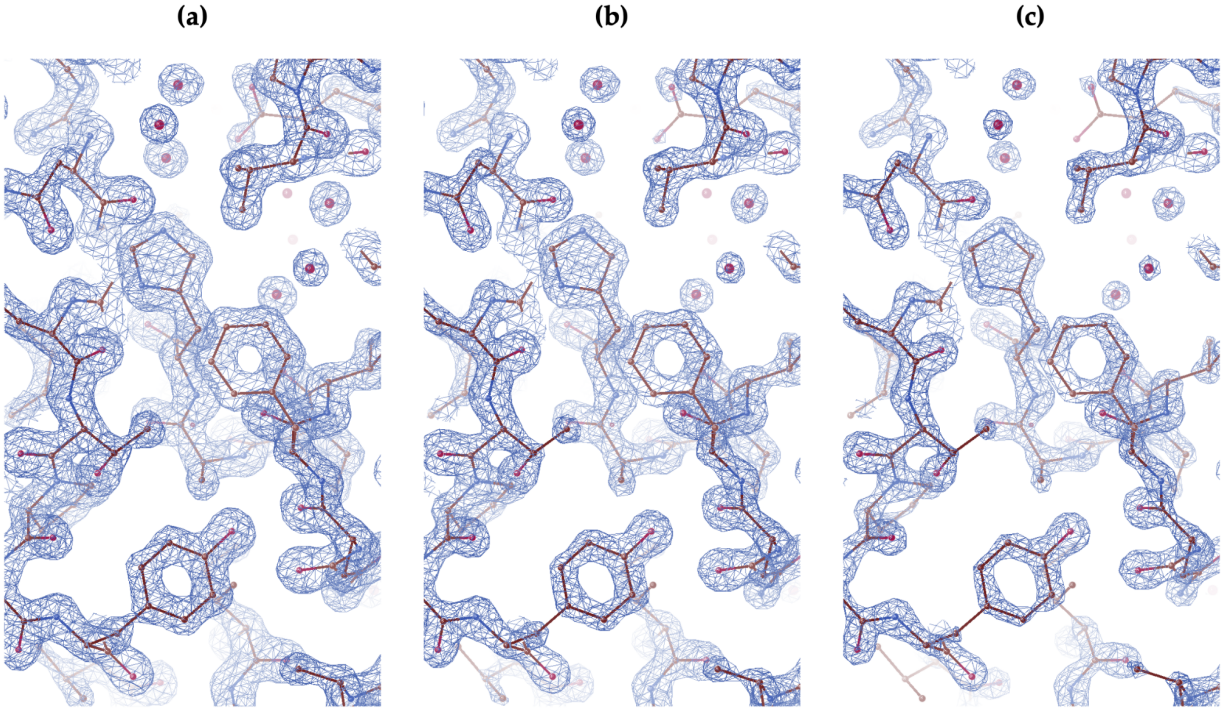

Figure 2. 2Fo-Fc map is colored in blue and contoured at (**a**) 2.0 σ, (**b**) 2.5 σ, and (**c**) 3.0 σ level. *Cb*FDH is shown in stick representation.

3.2. Two homodimers are nearly identical with minor conformational changes

We superposed the two *Cb*FDH homodimers in our crystal structure to compare their similarities. The two homodimers aligned with an RMSD score of 0.352. Particularly, the dimerization regions within the *Cb*FDH homodimers seem almost identical (Figure 3c) while some minor conformational changes are observed in regions farther from the central dimerization core domain (Figure 3b,d).

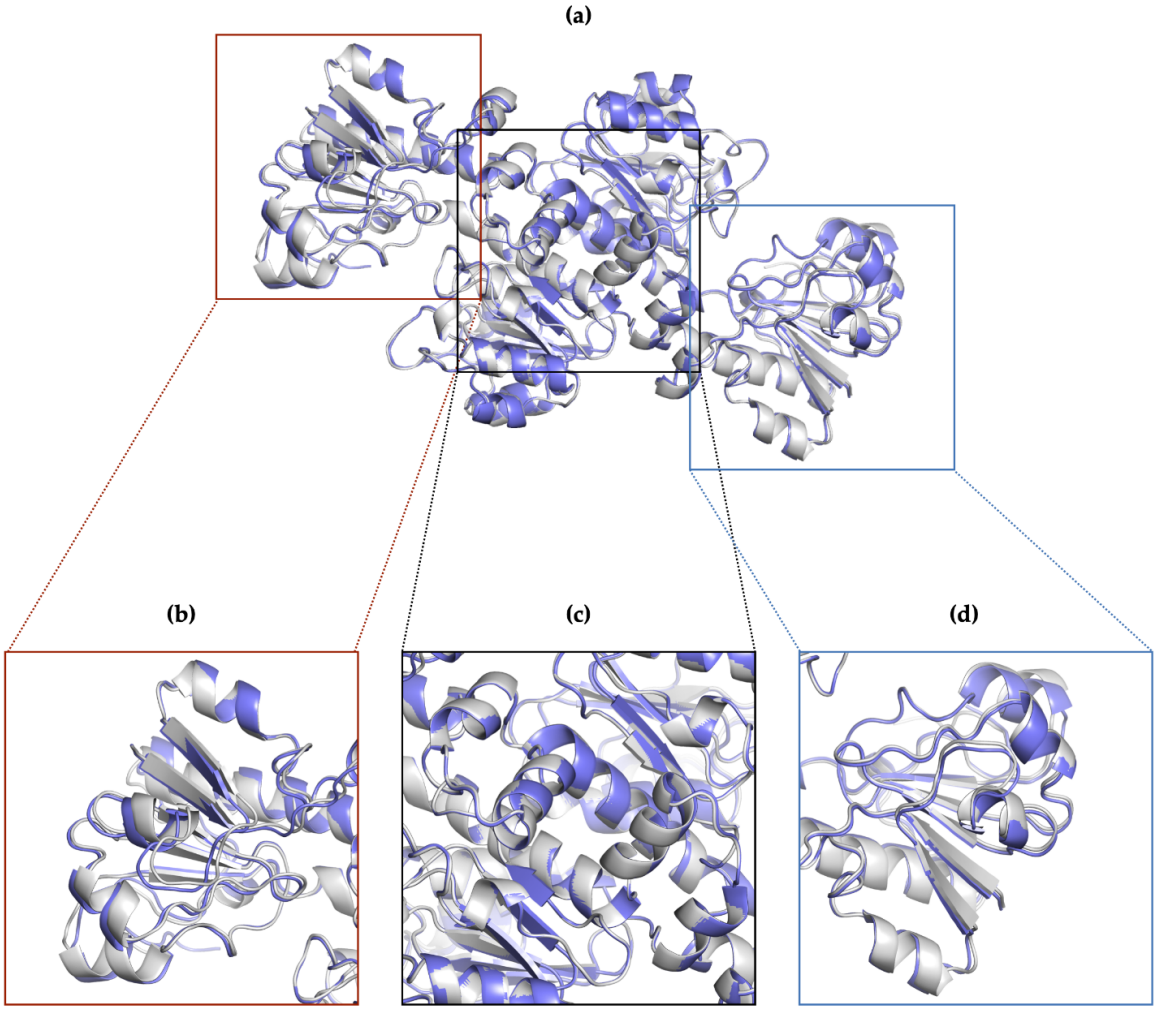

Figure 3. Superposition of two homodimers of *Cb*FDH. (**a**) Two homodimers superposed. The first homodimer is colored in slate, and the second one is colored in gray. (**b**),(**d**) Regions that do not participate in dimerization. (**c**) Dimerization region.

3.2. Two apo structures align with some minor differences

When we compare our apo 1.4 Å *Cb*FDH structure with a previously published apo structure (PDB ID: 5DNA) at 1.75 Å resolution, the overall structures align very well with an RMSD of 0.266 Å (Figure 4).

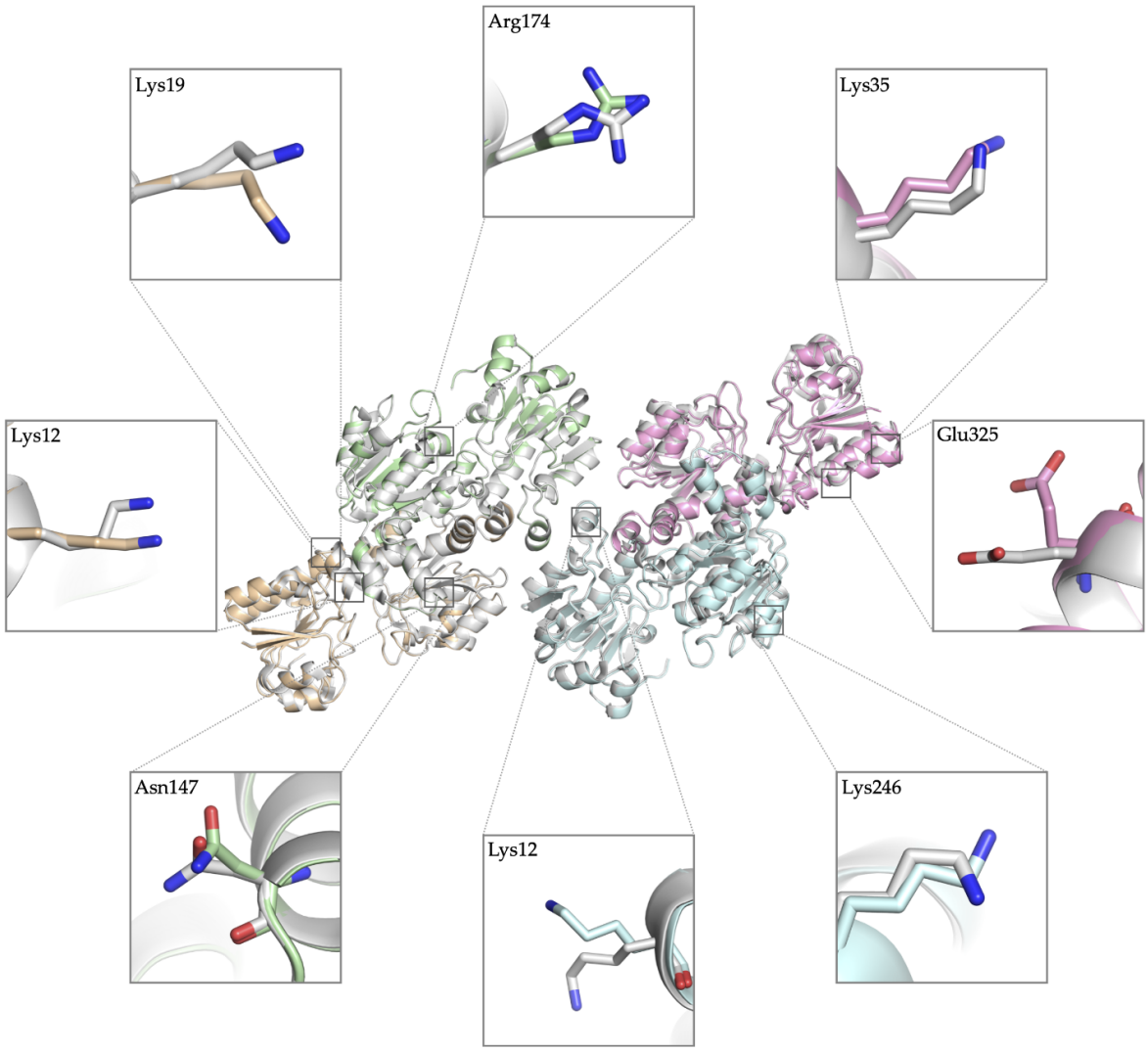

Figure 4. Superposition of two apo wild-type *Cb*FDH structures. Our 1.4 Å structure is shown in four distinct colors indicating each chain; A: pale green, B: wheat, C: pink, D: pale cyan. 5DNA is colored gray. Conformational differences in side chains are shown in boxes in stick representation.

Differences are mostly observed in the residues that directly exposed to the solvent (Figure 4). In addition, the previous apo *Cb*FDH structure lacks the residues 15-18 (Ala15, Asp16, Glu17, Glu18) in chain C, while our structure has a well-defined electron density for them (Figure 5).

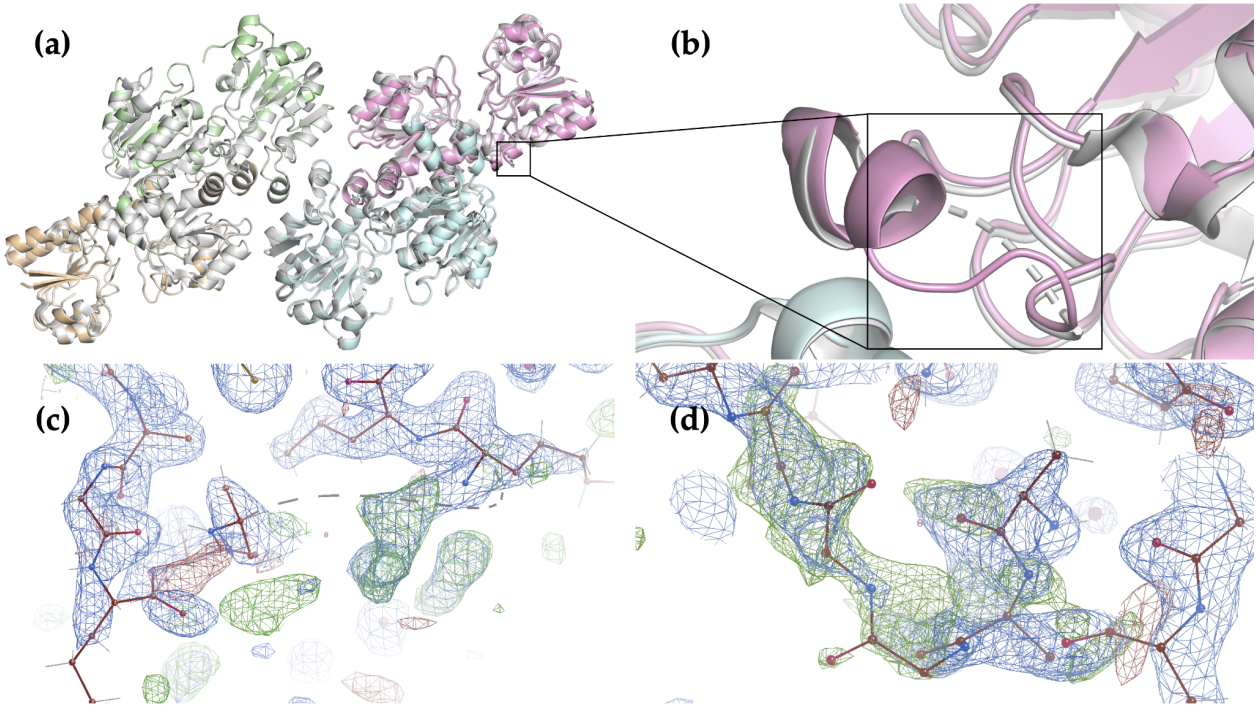

Figure 5. Chain C residues 15-18 comparison between our *Cb*FDH and 5DNA. (**a**) Superposition of our structure with 5DNA. The black square indicates the residues 15-18. Our structure is colored in four colors: pale green, wheat, pink, pale cyan. 5DNA is colored in gray. (**b**) A closer look at the residues 15-18 in the black square. The dashed lines represent the missing residues of 5DNA. (**c**) 2Fo-Fc (colored in slate) and Fo-Fc (colored in green) maps showing the 5DNA chain C residues 14 and 19 and dashed lines for the missing residues. (**d**) 2Fo-Fc and Fo-Fc maps colored same is panel C showing the chain C residues 14-19 belonging to our *Cb*FDH structure.

3.3. Comparison of apo and holo CbFDH structures reveal major conformational changes

We also compared our structure with a previously published holo *Cb*FDH structure (PDB ID: 5DN9) [15]. The holo structure contains the coenzyme NAD^+, a chloride ion, and the FDH inhibitor azide. Two structures superposed with an overall RMSD of 1.80 Å, indicating the presence of major conformational differences (Figure 6).

The difference was observed not only in the more flexible loop regions, but also in domainc containing multiple α-helices and β-sheets. Morph analysis was performed in order to reveal how the *Cb*FDH complex changes its conformation in the presence of both the coenzyme NAD^+ and inhibitor azide [3]. Our structure was used as the starting conformation and the holo structure was the end conformation. The morph analysis revealed that binding of NAD^+ and azide

induces/stabilizes major conformational changes. Upon binding, the two domains come closer together to close the catalytic site binding cleft (Supplementary Movie 1 & 2).

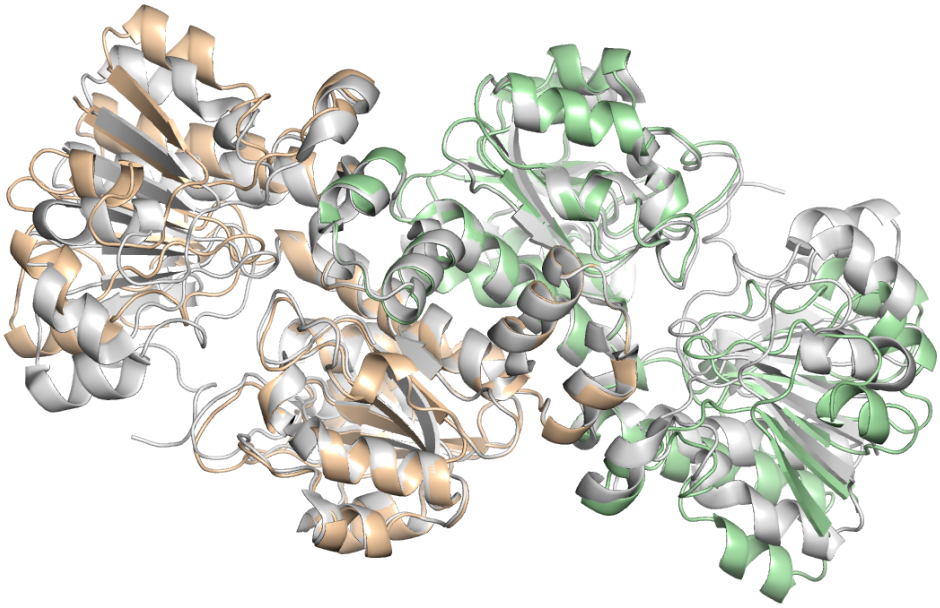

Figure 6. Superposition of apo and holo (5DN9) wild-type *Cb*FDH structures. The 1.4-Å apo structure is shown in two colors: wheat and pale green. The holo structure is shown in gray.

3.4. Thermal ellipsoids show the significant flexibility differences between apo and holo structures

We compared the flexibility of our structure and the holo structure (PDB ID: 5DN9) by examining the *B-factor* distribution of each. In our structure, the thermal ellipsoids are larger and more elongated, with warmer colors ranging from red to yellow to green (Figure 7a). In contrast, the ellipsoids in the holo structure (PDB ID: 5DN9) are smaller and more spherical, with mostly dark blue color (Figure 7b).

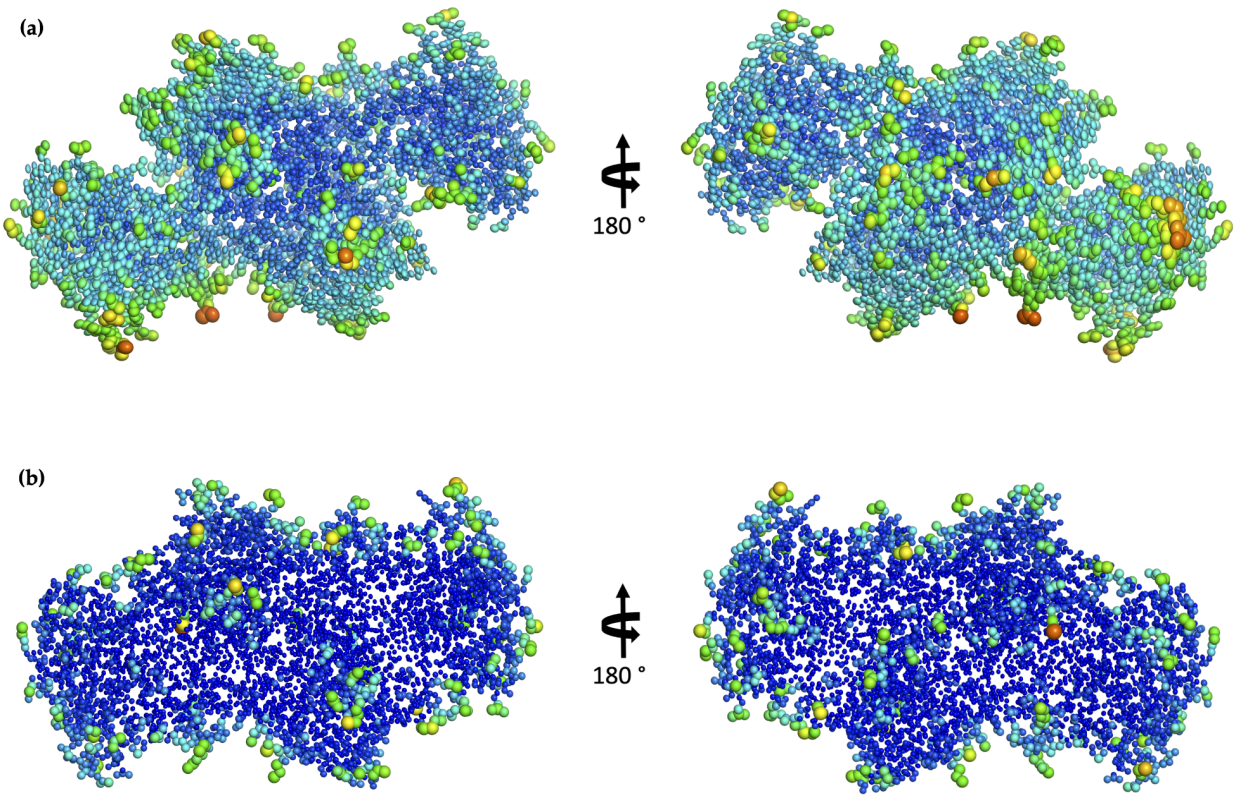

Figure 7. Thermal ellipsoid representation of (**a**) our apo *Cb*FDH structure and (**b**) holo structure (PDB ID: 5DN9).

4. Discussion

Our 1.4 Å-resolution apo *Cb*FDH structure provides structural details of the enzyme in its homodimeric form. The superior-quality electron density map reveals atomic details of the amino acid residues and water molecules. This high-resolution structure revealed a high degree of similarity between the two subunits, with only minor differences observed in the more flexible regions outside of the dimerization core domain region. These minor changes are expected since they can move more freely as they do not participate in dimer formation.

While the two homodimers in our structure were highly identical, a comparison with a previously published apo structure, consisting of two homodimers as well, (PDB ID: 5DNA) at a 1.75 Å resolution revealed several minor differences, particularly in the side chains of solvent exposed charged residues. Furthermore, our structure also showed improved electron density for flexible loop residues 15-18 (Ala15, Asp16, Glu17, Glu18) in chain C that were not well defined in the previous structure.

In addition to the apo structure, we also analyzed the conformational changes that occur in *Cb*FDH when it is bound by the coenzyme NAD^+ and the inhibitor azide. Azide is used as an inhibitor of *Cb*FDH due to having a linear, triatomic structure similar to carbon dioxide, the substrate of the reverse reaction [3]. When we superposed our structure with the holo form, we observe major conformational changes through the entire structure with a high RMSD value of 1.80 Å. We performed morphing between the two conformations to monitor how the structure changes globally and within the binding pocket. We observed significant conformational changes resulting in the closing of the binding pocket, which precluded the entry of another coenzyme or substrate. Also, this movement brings the catalytic domain and binding domain closer to a more compact state so that the reaction can take place.

Having observed significant conformational changes throughout the enzyme upon coenzyme and substrate/inhibitor binding, we further compared the apo and holo structures by examining the thermal ellipsoids. Thermal ellipsoids indicate the direction and magnitude of the thermal vibrations in crystal structures. Their size, shape, and color give information about the movement of the atoms within a structure. Our comparison of thermal ellipsoids revealed that the apo form is more flexible and dynamic compared to the holo form, suggesting that the enzyme becomes more stable and rigid upon binding of the coenzyme and substrate/inhibitor. In fact, it is already known that NAD^+ and inhibitor binding causes the binding pocket to close via global conformational changes, making the protein more compact.

The detailed analysis of the wild-type *Cb*FDH structure presented in this study provides comprehensive insights into this enzyme's structural dynamics, and highlights its potential for further time-resolved serial X-ray kineto-crystallography studies. The collective information gathered from these studies will be invaluable for future enzymology and protein engineering efforts to improve the enzymatic activity and stability for sustainable and efficient industrial applications.

Author Contributions: M.G., B.Y., H.B., and H.D. designed the experiment. M.G. and B.Y. performed the protein expression, purification, and crystallization. M.G. and B.Y. performed the data collection, processing, and structure refinement. M.G., B.Y., H.B., and H.D. prepared the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: H.D. acknowledges support from NSF Science and Technology Center grant NSF-1231306 (Biology with X-ray Lasers, BioXFEL). This publication has been produced benefiting from the 2232 International Fellowship for Outstanding Researchers Program of the TÜBİTAK (Project No. 118C270). However, the entire responsibility of the publication belongs to the authors of the publication. The financial support received from TÜBİTAK does not mean that the content of the publication is approved in a scientific sense by TÜBİTAK.

Data Availability Statement: The apo *Cb*FDH presented in this manuscript has been deposited to the Protein Data Bank under the accession number 8HTY. Any remaining information can be obtained from the corresponding author upon request.

Acknowledgments: Authors would like to dedicate this manuscript to the memory of Dr. Albert E. Dahlberg and Dr. Nizar Turker. The authors gratefully acknowledge use of the services and Turkish Light Source (*Turkish DeLight*) X-ray facility at Sağlık Bilimleri University Deneysel Tıp Araştırma ve Uygulama Merkezi (SBU-DETUAM). The authors also gratefully acknowledge use of the services and facilities at Koç University Isbank Research Centre for Infectious Diseases (KUIS-CID).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nielsen, C.F.; Lange, L.; Meyer, A.S. Classification and enzyme kinetics of formate dehydrogenases for biomanufacturing via CO_2 utilization. *Biotechnol Adv* **2019**, 37, 107408, doi:10.1016/j.biotechadv.2019.06.007.
2. Niks, D.; Hille, R. Reductive activation of CO_2 by formate dehydrogenases. *Methods Enzymol* **2018**, 613, 277-295, doi:10.1016/bs.mie.2018.10.013.
3. Popov, V.O.; Lamzin, V.S.; NAD^+-dependent formate dehydrogenase. *Biochem J* **1994**, 301, 625-643, doi:10.1042/bj3010625.
4. Bulut, H.; Yuksel, B.; Gul, M.; Eren, M.; Karatas, E.; Kara, N.; Yilmazer, B.; Kocyigit, A.; Labrou, N.E.; Binay, B. Conserved Amino Acid Residues that Affect Structural Stability of *Candida boidinii* Formate Dehydrogenase. *Appl Biochem Biotechnol* **2021**, 193, 363-376, doi:10.1007/s12010-020-03429-0.
5. Murashima, K.; Yoneda, H.; Sumi, H.; Amao, Y. Electrocatalytic production of formaldehyde with formaldehyde dehydrogenase using a viologen redox mediator. *New J Chem* **2022**, 46, 10004-10011, doi:10.1039/d2nj00692h.
6. Powers, P.O. Phenol-, Urea-, and Melamine-Formaldehyde Plastics. *Ind Eng Chem* **2002**, 45, 1063-1066, doi:10.1021/ie50521a053.

7. Formaldehyde. In *Meyler's Side Effects of Drugs*, 16th ed.; Aronson, J.K., Ed.; Elsevier: 2016; pp. 437-443, doi:10.1016/B978-0-444-53717-1.00776-9.
8. Schuler, E.; Demetriou, M.; Shiju, N.R.; Gruter, G.M. Towards Sustainable Oxalic Acid from CO_2 and Biomass. *ChemSusChem* **2021**, 14, 3636-3664, doi:10.1002/cssc.202101272.
9. Eppinger, J.; Huang, K.-W. Formic Acid as a Hydrogen Energy Carrier. *ACS Energy Letters* **2016**, 2, 188-195, doi:10.1021/acsenergylett.6b00574.
10. Garman, E.F.; Owen, R.L. Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2006**, 62, 32-47, doi:10.1107/S0907444905034207.
11. Atalay, N.; Akcan, E.K.; Gul, M.; Ayan, E.; Destan, E.; Ertem, F.B.; Tokay, N.; Çakilkaya, B.; Nergiz, Z.; Karakadioğlu, G.; et al. Cryogenic X-ray crystallographic studies of biomacromolecules at Turkish Light Source "Turkish DeLight". *bioRxiv* **2022**, 2022.2009.2003.506456, doi:10.1101/2022.09.03.506456.
12. Rigaku. CrysAlisPro Software System, Version 1.171.42.35a. **2021**, Rigaku Oxford Diffraction, https://www.rigaku.com.
13. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J Appl Crystallogr* **2007**, 40, 658-674, doi:10.1107/S0021889807021206.
14. Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010**, 66, 213-221, doi:10.1107/S0907444909052925.
15. Guo, Q.; Gakhar, L.; Wickersham, K.; Francis, K.; Vardi-Kilshtain, A.; Major, D.T.; Cheatum, C.M.; Kohen, A. Structural and Kinetic Studies of Formate Dehydrogenase from *Candida boidinii. Biochemistry* **2016**, 55, 2760-2771, doi:10.1021/acs.biochem.6b00181.
16. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, 60, 2126-2132, doi:10.1107/S0907444904019158.
17. The *PyMOL* Molecular Graphics System, Version 2.5.2, Schrödinger, LLC.
18. Gul, M.; Ayan, E.; Destan, E.; Johnson, J.A.; Shafiei, A.; Kepceoglu, A.; Yilmaz, M.; Ertem, F.B.; Yapici, I.; Tosun, B.; et al. Rapid and High Resolution Ambient Temperature Structure Determination at Turkish Light Source. *bioRxiv* **2022**, 2022.10.12.511637, doi:10.1101/2022.10.12.511637.