

Ni(II)-binding affinity of CcNikZ-II and its homologs

Diep, Patrick; Stogios, Peter J; Evdokimova, Elena; Savchenko, Alexei; Mahadevan, Radhakrishnan; Yakunin, Alexander F

Febs Journal

DOI:
[10.1111/febs.17125](https://doi.org/10.1111/febs.17125)

E-pub ahead of print: 31/03/2024

Publisher's PDF, also known as Version of record

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Diep, P., Stogios, P. J., Evdokimova, E., Savchenko, A., Mahadevan, R., & Yakunin, A. F. (2024). Ni(II)-binding affinity of CcNikZ-II and its homologs: the role of the HH-prong and variable loop revealed by structural and mutational studies. *Febs Journal*. Advance online publication. <https://doi.org/10.1111/febs.17125>

Hawliau Cyffredinol / General rights



Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Ni(II)-binding affinity of CcNikZ-II and its homologs: the role of the HH-prong and variable loop revealed by structural and mutational studies

Patrick Diep^{1,2} , Peter J. Stogios¹, Elena Evdokimova¹, Alexei Savchenko^{1,3}, Radhakrishnan Mahadevan^{1,4} and Alexander F. Yakunin^{1,5} 

¹ Department of Chemical Engineering and Applied Chemistry, BioZone – Centre for Applied Bioscience and Bioengineering, University of Toronto, Toronto, Ontario, Canada

² Systems & Synthetic Biology Group, Biosciences and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA

³ Department of Microbiology, Immunology and Infectious Diseases, University of Calgary, Calgary, Alberta, Canada

⁴ Institute of Biomedical Engineering, University of Toronto, Toronto, Ontario, Canada

⁵ Centre for Environmental Biotechnology, School of Natural Sciences, Bangor University, Bangor, Wales, UK

Keywords

crystallography; fluorescence quenching; metal-binding activity; metalloproteins; mutagenesis

Correspondence

P. Diep and A. F. Yakunin, Department of Chemical Engineering and Applied Chemistry, BioZone – Centre for Applied Bioscience and Bioengineering, University of Toronto, Toronto, ON, Canada
 Tel: +1 (925) 960 3259
 E-mail: diep5@lnl.gov (PD); a.yakounine@bangor.ac.uk (AFY)

Extracytoplasmic Ni(II)-binding proteins (NiBPs) are molecular shuttles involved in cellular nickel uptake. Here, we determined the crystal structure of apo CcNikZ-II at 2.38 Å, which revealed a Ni(II)-binding site comprised of the double His (HH-)prong (His511, His512) and a short variable (v-) loop nearby (Thr59-Thr64, TEDKYT). Mutagenesis of the site identified Glu60 and His511 as critical for high affinity Ni(II)-binding. Phylogenetic analysis showed 15 protein clusters with two groups containing the HH-prong. Metal-binding assays with 11 purified NiBPs containing this feature yielded higher Ni(II)-binding affinities. Replacement of the wild type v-loop with those from other NiBPs improved the affinity by up to an order of magnitude. This work provides molecular insights into the determinants for Ni(II) affinity and paves way for NiBP engineering.

(Received 13 November 2023, revised 30 January 2024, accepted 13 March 2024)

doi:10.1111/febs.17125

Introduction

Metal ions (metals) are an indispensable part of all cells due to their unique properties that enable more complex biochemistry and regulation [1,2]. Almost half of known enzymes and proteins require or associate with various metals to function. In particular, several transition metals of the first-row d-block elements like Cu(II), Fe(II, III), Co(II), and Ni(II) (nickel) are

indispensable cofactors for the activity of many enzymes. These metals are thus critical for bacterial adaptation to challenging environments such as pathogenesis in the anaerobic gastrointestinal tract of humans, resistance to extremely acidic mine waters, and survival in metal-scarce soil microenvironments [3–5]. The primary high-affinity metal uptake systems

Abbreviations

AB, activity buffer; ABC, ATP-binding cassette; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; ICP-MS, inductively coupled plasma mass spectrometry; ITC, isothermal titration calorimetry; ITFQ-LTA, intrinsic tryptophan fluorescence quenching (ligand titration assay); LB, Luria-Bertani (media); MWCO, molecular weight cut-off; NiBP, nickel (II)-binding protein; NTA, nitriloacetic acid; PDB, Protein Data Bank; PTFE, polytetrafluoroethylene; SBP, solute-binding protein; SDS/PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TB, terrific broth (media); TEV, Tobacco Etch Virus.

are represented by ATP-binding cassette (ABC)-type transporters, which include the canonical ABC-type importers composed of two channel-forming transmembrane proteins, two nucleotide-binding proteins, and an extracytoplasmic solute-binding protein (SBP) [6–8]. The periplasmic and extracellular SBPs of ABC-type transporters function as molecular tools for binding substrates (metals, peptides) and are the main determinants of the transporter selectivity. The large and diverse SBP superfamily includes over 147 000 proteins, which have relatively low sequence similarity, but share a similar overall three-dimensional (3D) structure [9,10]. These proteins are capable of binding very different substrates with Ni(II) specific proteins (NiBPs) associated with peptide ABC transporters: PepT family in the transporter classification or Cluster C in the structural SBP classification [9–12]. Compared to other SBPs, the Cluster C proteins are larger (55–70 kDa) due to the presence of an additional domain, but based on sequence they can not be divided into subgroups according to substrate selectivity (e.g., peptides, arginine, cellobiose, or nickel ions) [13].

Currently, there are over 750 3D structures of unique SBPs in the Protein Database (PDB) with the sub-cluster C-I comprising 14 metal-specific SBP structures [9,10,14]. However, little functional information about NiBPs is currently available as only several proteins have been biochemically and structurally characterized, these include CjNikZ from *Campylobacter jejuni* [13,15], BsNikA from *Brucella suis* [13], YpYntA from *Yersinia pestis* [13], SaNikA and SaCntA from *Staphylococcus aureus* [16], HhNikA from *Helicobacter hepaticus* [17], HpCeuE from *Helicobacter pylori* [18], VpNikA from *Vibrio parahaemolyticus* [19], and CdOppA from *Clostridium difficile* [20]. Metalloproteins and SBPs like those listed are involved in natural metal homeostasis systems, but have also been subjects of engineering efforts for metal removal and recovery applications [21–26]. In natural proteins, metal preference and binding affinity for different divalent metal ions are expected to follow the Irving–Williams series of empirical rules (from weak to tight: Mg(II) < Mn(II) < Fe(II) < Co(II) < Ni(II) < Cu(II) > Zn(II)) [27]. Nevertheless, bacterial cells and proteins are able to overcome these metal-binding preferences using specific metal delivery systems, complex homeostatic machinery, and selective metal coordination sites, based on a precise arrangement of protein residues [1,28–30]. To re-design these proteins for altered binding to Ni(II) and other transition metals, a basic understanding of the molecular details underlying their affinity is needed, but remains poorly understood [6,13].

In our prior work, we explored the scalable use of microplate-based ITFQ for the analysis of Ni(II)-binding affinity of CjNikZ and CcNikZ-II from *Clostridium carboxidivorans*, which is the second NiBP in the *C. carboxidivorans* acetogenesis operon [31]. We demonstrated that both CcNikZ-II and CjNikZ exhibit micromolar affinities for Ni(II). To provide insights into the molecular details underlying the affinity of CcNikZ-II and CjNikZ to Ni(II), we determined the crystal structure of apo CcNikZ-II by X-ray crystallography at 2.38 Å and compared the ALPHAFOLD2-predicted Ni(II)-binding sites of nine NiBPs homologous to CcNikZ-II [32,33]. By combining these structural insights with metal-binding assays, alanine scanning mutagenesis, phylogenetic analyses, and binding site engineering, we characterized the role of the CcNikZ-II residues in Ni(II)-binding affinity and engineered higher affinity variants. These results provide new insights into the metal binding determinants of NiBPs, which will enable efforts to further improve their binding affinity and potentially their selectivity for different metals.

Results and Discussion

Crystal structure and alanine scanning mutagenesis of CcNikZ-II

CcNikZ-II was recombinantly expressed in *Escherichia coli* with the endogenous signal peptide (30 aa) replaced by a His₆-tag and affinity purified to over 95% homogeneity (Materials and methods, Fig. S1). Purified CcNikZ-II was crystallized using the hanging drop method, and the crystal structure of apo CcNikZ-II was determined at 2.38 Å resolution by molecular replacement (Table S1, PDB: 8EFZ). We attempted to crystallize the CcNikZ-II in its Ni-bound holo forms, but crystals only grew for the apoprotein. Preliminary analytical size exclusion chromatography revealed CcNikZ-II exists as a monomer (data not shown). Like other SBPs of Family 5, the CcNikZ-II monomer has a teardrop-like shape with two α/β domains: domain I with two subdomains (Ia: residues 1–40, 154–243, 463–526; Ib: 41–153) and domain II (244–462), which encloses a metal-binding site located between these lobes (Fig. 1A). The CcNikZ-II domain II is connected to subdomain Ia through a shared anti-parallel pleated β-sheet acting as a rigid hinge based on the apo and holo structure of CjNikZ (PDB: 4OET, 4OEV). No disulfide bonds were detected in the structure. A Dali search [34] for structurally homologous proteins in the PDB identified over one hundred homologous α/β fold proteins with low

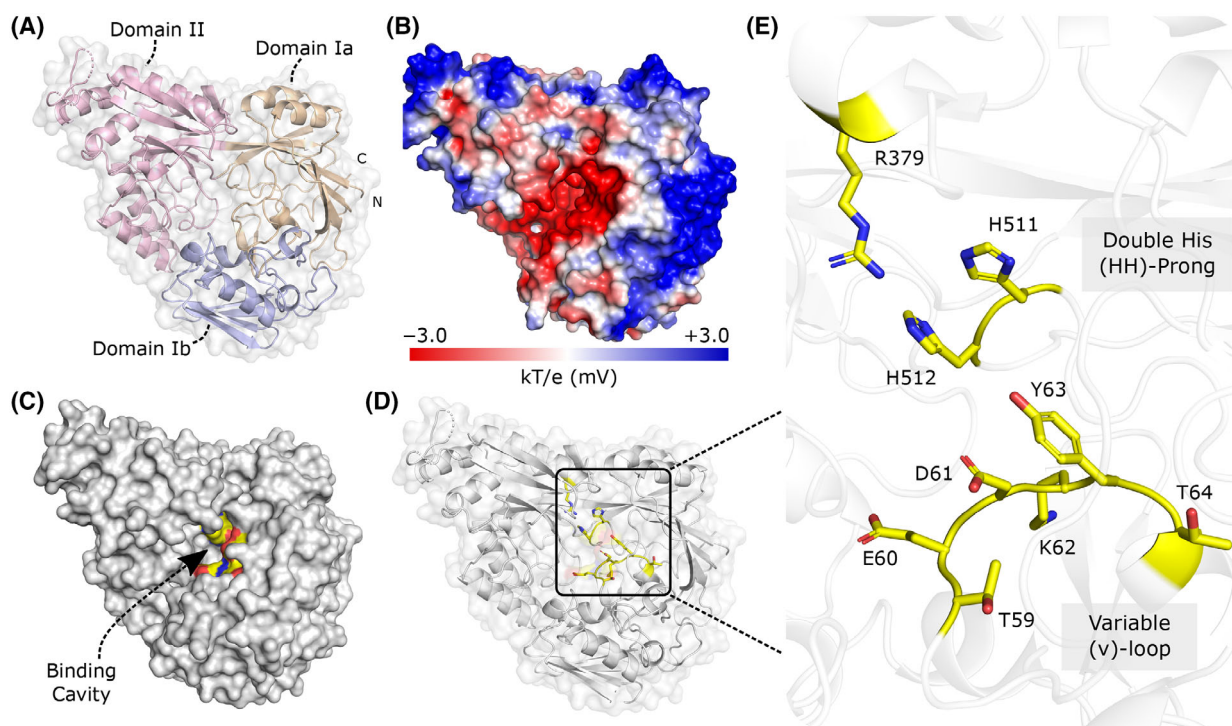


Fig. 1. Crystal structure of CcNikZ-II. (A) Overall fold of the CcNikZ-II protomer. The protein is comprised of two domains: the classic bi-lobal domain Ia (tan) and Ib (lavender), and domain II (rose). (B) An electrostatic surface representation generated using an adaptive Poisson-Boltzmann solver. Low kT/e values (red) demark negatively charged areas, and high kT/e values (blue) demark positively charged areas. (C) A surface representation of CcNikZ-II. The binding residues (yellow) are buried deep inside in the core of the protein. (D) Location of the Ni(II)-binding residues (yellow) with a (E) close-up view of the CcNikZ-II binding site. The Ni(II)-binding site is positioned between the three protein lobes with each lobe contributing residues toward metal binding. Red and blue atoms represent oxygen and nitrogen, respectively (C–E).

overall sequence similarity to CcNikZ-II (17–40% sequence identity). The top structural homologs of CcNikZ-II include CjNikZ from *Campylobacter jejuni* (PDB: 4OET, Z-score 51.2, rmsd 1.4 Å, 40% identity), the oligopeptide binding protein AppA from *Bacillus subtilis* (PDB: 1XOC, Z-score 42.2, rmsd 2.6 Å, 34% identity), and YpYntA from *Yersinia pestis* (PDB: 4OFL, Z-score 40.6, rmsd 2.5 Å, 25% identity). The electrostatic surface charge of CcNikZ-II revealed that the central cavity was negatively charged, which strongly suggested the presence of a binding site for positively charged metal ions like Ni(II). Immediately outside the cavity, a ring of neutral surface charge defined a perimeter separating the predominantly positively charged exterior from the cavity (Fig. 1B), which we suspect could be a mechanism for limiting adventitious surface binding of Ni(II) via repulsive forces, thus guiding the ions into the cavity.

Our recent study of purified CcNikZ and CjNikZ demonstrated that these proteins bind Ni(II) with micromolar affinities without the requirement for adding nickelophores such as histidine [31]. The structures of CjNikZ complexed with Ni(II) revealed His26,

His480, and His481 as direct Ni(II) ligands, whereas the remaining three ligands to complete the octahedral coordination geometry were provided by a bound free His (PDB: 4OEU) or bound oxalate (and a water molecule, PDB: 4OEV) acting as additional Ni(II) chelators [13]. The CjNikZ Ni(II)-binding site also included Arg344 coordinating the bound Ni-chelator (free His or oxalate), as well as several aromatic residues positioned near the bound Ni(II) (Tyr374, Phe380, Trp384, and Phe413). These insights from CjNikZ informed our analysis of the apo CcNikZ-II structure (Fig. 1C), which revealed the presence of several conserved residues buried within the protein metal-binding cavity (Fig. 1D,E). Similar to His480 and His481 in CjNikZ, the CcNikZ-II His511 and His512 (HH-prong) are positioned such that their imidazole groups pointed into the central cavity located between the protein domains. In the CcNikZ-II structure, the presence of a chloride ion (not shown) coordinated by the guanidino group of the conserved Arg379 suggested that like Arg344 in CjNikZ, this residue can be involved in Ni(II) coordination via a protein bound water molecule or a salt bridge with a potential small chelator

(e.g., oxalate or histidine). Superimposition of the apo structures of CcNikZ-II and CjNikZ (PDB: 4OET) revealed the structural conservation of these Ni(II)-binding residues (Arg379, His511, and His512 in CcNikZ-II and Arg344, His480, and His481 in CjNikZ) (Fig. 2A,B). The other highly conserved residues of the CcNikZ-II nickel-binding site include Asp376, Phe415, Trp419, and His428. However, the third His residue of CjNikZ (His26) directly interacting with the bound Ni(II) ion and located on a flexible loop below the HH-prong was missing in CcNikZ-II. Furthermore, the CjNikZ flexible loop contained one additional residue (A28) in the middle compared to CcNikZ-II (Fig. 2A,B). Taken together, the CcNikZ-II structure reveals the presence of a Ni(II)-binding site with several conserved residues, but also some differences compared to CjNikZ suggesting that there may be variations in how the two proteins bind to Ni(II).

The variable flexible loop (v-loop) appears to play a dynamic role in Ni(II)-binding based on the apo and holo CjNikZ structures (PDB: 4OET, 4OEU, 4OEV) (Fig. 2A,B). In CjNikZ, the v-loop undergoes a large

(> 5 Å) movement to reposition its residues (specifically His26) closer to the HH-prong for direct Ni(II)-binding [13]. The absence of His in the CcNikZ-II v-loop prompted us to use alanine scanning mutagenesis to study the impact of binding site residues on the Ni(II)-binding affinity of this protein (Fig. 3, Table S2, Fig. S2). We initially targeted residues in the first coordination sphere: Arg379, His511, His512, as well as generated a double Ala substitution of the HH-prong (H511A, H512A). Since Lys62 of the CcNikZ-II v-loop (TEDKYT) was structurally aligned with the CjNikZ-II v-loop His26, we hypothesized it had similar functionality and included it in the initial round of mutagenesis (K62A). Ala substitution of His511 and His512 each lead to a 2.3- and 1.2-fold decrease in binding affinity, respectively, implying that His511 plays a larger role in Ni(II)-binding than His512 (Fig. 3). Surprisingly, the mutant protein K62A had a 4.3-fold improvement in Ni(II)-binding affinity (Fig. 3). This was likely due to the positive charge of the Lys62 side chain amino group at the assay pH (7.2) that would cause it to repel Ni(II), and its

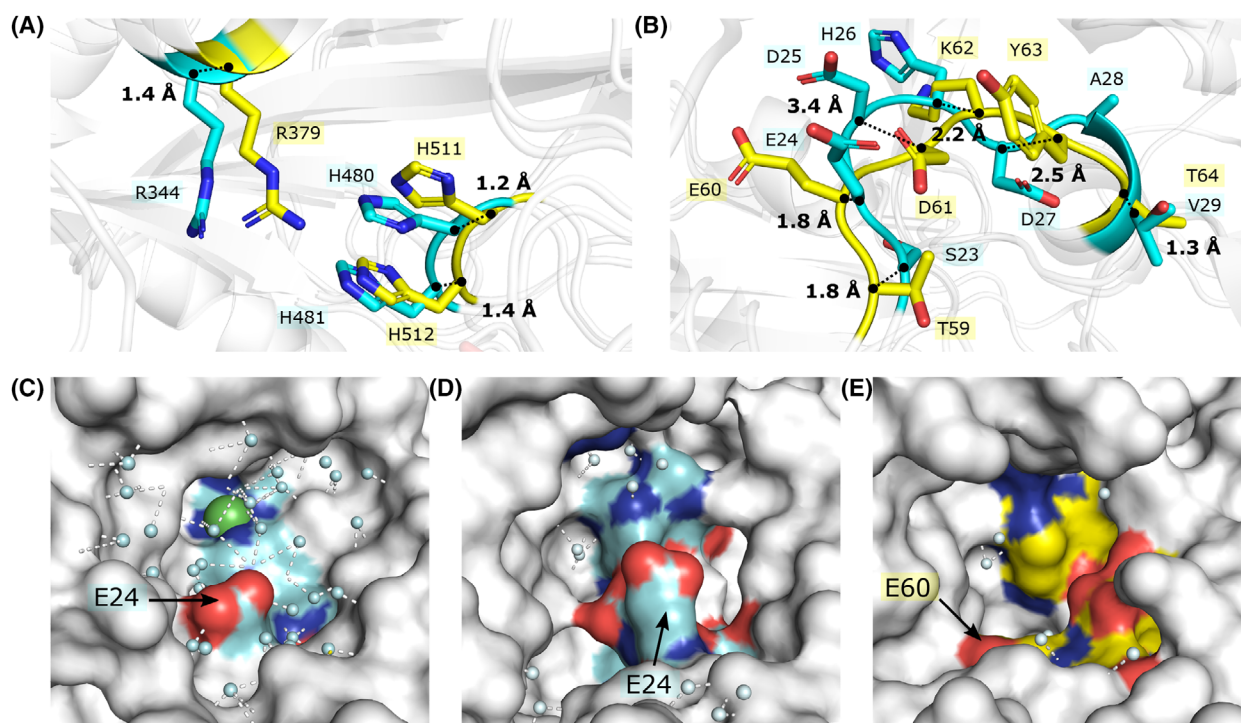


Fig. 2. Structural comparison of CcNikZ-II and CjNikZ. (A, B) Close-up view of Ni(II)-binding site: structural superimposition of apo CcNikZ-II (yellow) and apo CjNikZ (PDB: 4OET, cyan) is shown for the upper half (A) and the bottom half (B) of the metal-binding site. Black dotted lines represent the distance between the α carbons for each aligned residue. (C–E) Close-up view of the surface of Ni(II)-bound holo CjNikZ-II (PDB: 4OEV) (C), apo CjNikZ (PDB: 4OET) (D), and apo CcNikZ-II (E). Water molecules (blue spheres) and their potential polar contacts (white dashed lines) are displayed to show the absence and presence of water networks formed within the binding site. Nickel (green sphere) is shown buried in the site behind the water network. Red and blue atoms represent oxygen and nitrogen, respectively.

bulkiness that may render the v-loop less mobile, which may reflect a preference for another substrate (e.g., naked ion vs. complexed ions). The R379A and double H511A/H512A mutant proteins were found to be poorly expressed in *E. coli* and could not be purified. Thus, these Ala substitution experiments confirmed the important role of the conserved residues His511, His512, and Lys62 in Ni(II)-binding by CcNikZ-II. The increased Ni(II)-binding affinity of the CcNikZ-II K62A mutant protein led us to probe the role of other residues within the v-loop in metal binding. The structure of CjNikZ complexed with nickel (PDB: 4OEV) showed that the side chain of Glu24 of its v-loop protruded directly into the water network formed in the nickel-binding site (Fig. 2C) [13]. In contrast, a water network was not observed in the Ni(II)-binding cavities of apo CjNikZ (PDB: 4OET) and apo CcNikZ-II structures (Fig. 2D,E). We therefore hypothesized that in CcNikZ-II the v-loop Glu60 (TEDKYT) also plays a critical role in Ni(II)-binding by participating in the hydrogen-bonding network.

In the second round of mutagenesis (Fig. 3), we substituted each residue of the v-loop sequence TEDKYT with alanine (which also included K62A from the first round). Ni(II)-binding assays confirmed our hypothesis regarding Glu60 by showing a large (8.4-fold) decrease in Ni(II)-binding affinity of purified E60A mutant protein (Fig. 3). This was further substantiated with the full substitution of the CcNikZ-II v-loop by alanines (TEDKYT>AAAAAA) that revealed a 3.6-fold decrease in Ni(II)-binding affinity

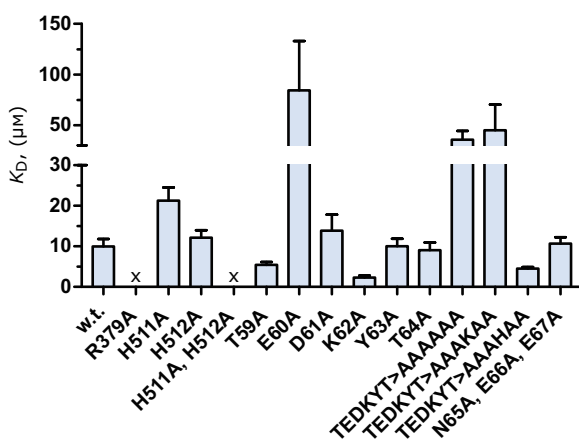


Fig. 3. Ni(II)-binding affinity of the wild type and mutant CcNikZ-II proteins. Ni(II)-binding affinity (K_D) of purified wild type (w.t.) CcNikZ-II and various mutant proteins (single, double, and triple Ala-substitutions and v-loop substitutions). 'x' denotes mutant proteins that were poorly expressed and could not be purified. Experimental triplicates were performed ($n=3$). Errors bars: SD.

(Fig. 3). Similarly, substituting out all residues in the TEDKYT v-loop to alanine, except Lys62 (TEDKYT>AAAKAA) produced a mutant protein which behaved like the TEDKYT>AAAAAA protein (Fig. 3). Remarkably, replacing the lysine in this mutant protein with a histidine (TEDKYT>AAAHAA) improved the binding affinity 2.2-fold over the wild-type CcNikZ-II (TEDKYT). The C-terminal end of the CcNikZ-II v-loop is connected to a small α -helix with the sequence NEE. Given its spatial distance from H511 and H512, we expected no change in Ni(II)-binding affinity upon alanine substitution, which we confirmed with the triple mutant protein (N65A, E66A, E67A) that showed little change in affinity (Fig. 3). Together, this pointed to the necessity of Glu60 and the advantage of having a closely adjacent histidine suggesting that core v-loops containing both a Glu and His within close proximity near the front of the loop would confer a higher Ni(II)-binding affinity to CcNikZ-II and other NiBPs. Milder changes in affinity observed in the D61A (1.3-fold decrease in affinity) and T59A (1.8-fold increase in affinity) mutant proteins were alone insufficient for elucidating their role. We also noted that the Y63A and T64A mutant proteins showed little change in affinity, suggesting residues in these positions may contribute less to Ni(II)-binding affinity. To further understand how v-loops can impact Ni(II)-binding affinity, we next used a phylogenetic approach to analyze the diversity of v-loops found in the NiBP family of CcNikZ-II and CjNikZ (Fig. 4, Cluster-12).

Phylogenetic analysis of NiBPs

A previous phylogenetic analysis of NiBPs using eight experimentally studied NiBPs (*EcNikA*, *YpYntA*, *VpNikA*, *HhNikA*, *CjNikZ*, *SaNikA*, *SaCntA*, *BsNikA*) as seeds revealed high sequence similarity between NiBPs and PepBPs suggesting a common evolutionary origin [13]. However, this approach introduced bias due to selection of the pool of sequences of experimentally characterized NiBPs suggesting that alternative modes of substrate-binding in NiBPs and PepBPs that could have promiscuity for Ni(II)-binding may have been under-represented and overlooked. Therefore, we performed a global phylogenetic analysis of members from both groups using InterPro where the SBPs are organized according to the Tam-Saier Classification (Fig. 4). We first decreased the redundancy of the InterPro family IPR030678 using the CD-HIT algorithm to obtain a representative library of 4561 NiBP and PepBP sequences [35,36]. We then used FastTree to generate a phylogenetic tree from a MAFFT alignment to obtain a representative library [37]. Therefore, we

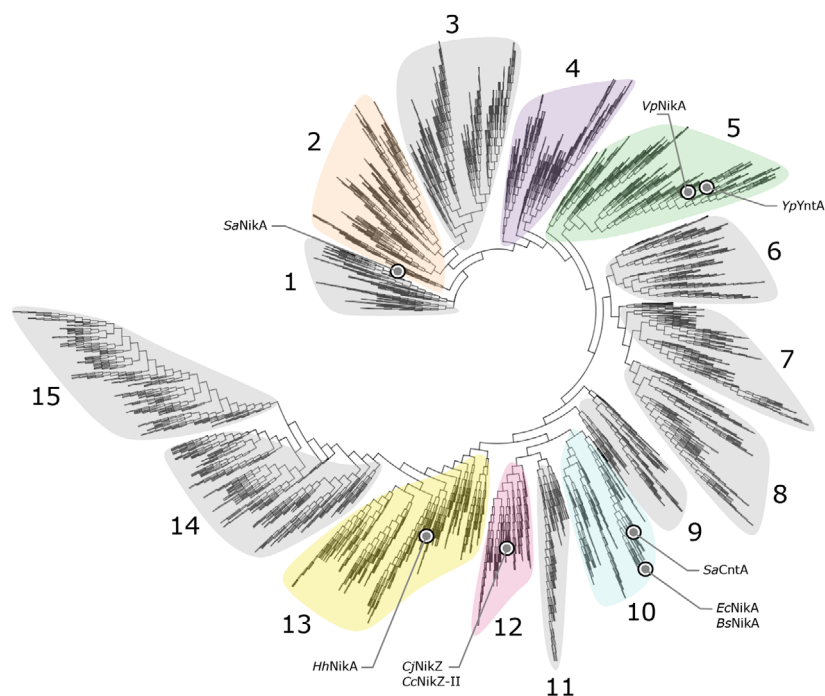


Fig. 4. Phylogenetic analysis of NiBPs. The representative library was created by applying the CD-HIT algorithm to the InterPro family IPR030678 such that sequences with 60% redundancy were removed. From this, the phylogenetic tree representation of the representative library built using FastTree on Geneious is shown. Experimentally characterized NiBPs (gray circles) are mapped onto the tree with colored spheres demarking their phylogenetic branch.

captured a fuller diversity of NiBP/PepBPs, which include at least 13 protein clusters.

Similar to Lebrette *et al.* [13], we found that *VpNikA* and *YpYntA* are positioned together in Cluster-5 while *SaCntA*, *EcNikA*, and *BsNikA* are grouped in Cluster-10 (Fig. 4). In contrast, *SaNikA* (Cluster-2) is positioned further away from *VpNikA* and *YpYntA* (Cluster-5) and closer to the root (Fig. 4). According to Lebrette *et al.* [13], *SaNikA* follows the third Ni(II)-binding strategy where the binding site residues interact with bound nickelophores (two free histidine amino acids) with no direct contact between the protein and nickel ion. The *SaNikA* metal-binding site also lacks the His residues found in other NiBPs that follow the outlined first and second binding strategies with direct contacts with bound Ni(II). Based on the positioning of *SaNikA* in the SBP phylogenetic tree, this suggests that ancestral proteins of the IPR030678 family may have been PepBPs originally, but because peptides are strong complexing agents for nickel, some members of the family may have evolved selectivity for nickel and nickel complexes. Furthermore, *CjNikZ* and *CcNikZ-II* (Cluster-12) appeared to be closer to *EcNikA* homologs (Cluster-10) than *VpNikA* and *YpYntA* (Cluster-5) in Lebrette *et al.*, suggesting that *CjNikZ* and *CcNikZ-II* shared a common ancestor with *EcNikA* despite having only 25% identity with each other.

In this study, we focused on SBPs containing the double His (HH) prong motif from Cluster-12, where

both *CcNikZ-II* and *CjNikZ* are located (close homologs) (Fig. 4). Additionally, by searching for the HH-prong motif across the SBP phylogenetic tree, we found homologous proteins in Cluster-4 (distant homologs) containing this motif, which was absent in other clusters. Upon closer inspection, 88% (21/24) of proteins from Cluster-12 and 10% (8/78) of proteins from Cluster-4 contained the C-terminal HH-prong. For comparative sequence analysis, we extracted sequences of nine close homologs (Cluster-12) and three distant homologs (Cluster-4) based on their uniqueness within their local phylogenies. Alignment of these sequences with *CcNikZ-II* and *CjNikZ* revealed three observations (Fig. 5A): (a) the positions of the arginine (Arg379 in *CcNikZ-II*) and HH-prong were highly conserved, (b) the Cluster-4 homologs had distinctly different sequences of the v-loop region, and (c) the Cluster-4 SBP protein from *Nocardiosis gilva* (*NgNikZ*) and Cluster-12 sequences had more similar v-loops, but the exact positional conservation was unclear. We therefore selected the sequences of 10 close and one distant (*NgNikZ*) homologs for comparative biochemical analysis of v-loop sequences and Ni(II)-binding affinity (Table S3, Fig. 5B).

Comparative analysis of v-loop sequences and structures

In the *CcNikZ-II* sequence, the v-loop section is flanked by two strongly conserved consensus

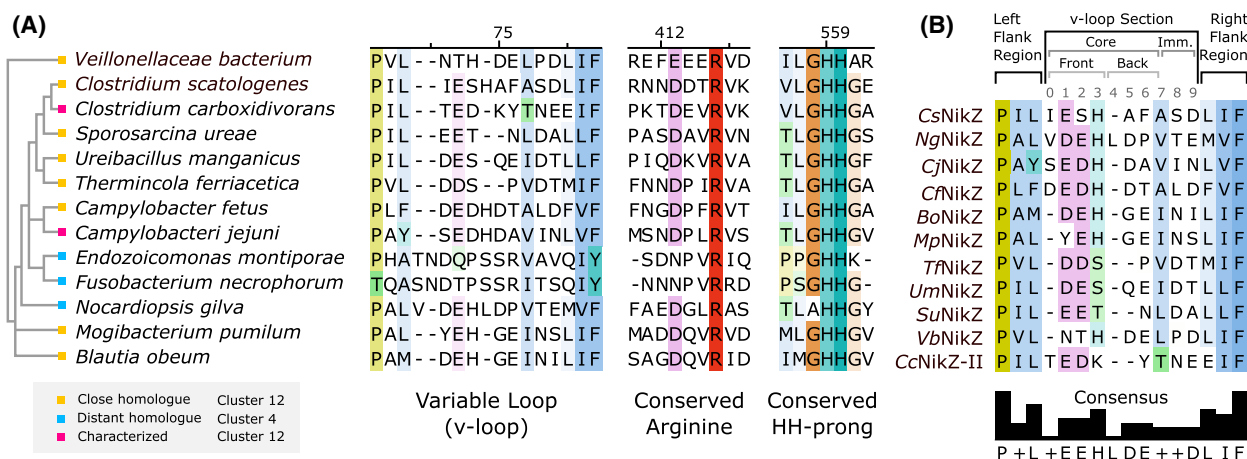


Fig. 5. The sequence motifs of selected NiBPs involved in Ni(II)-binding. (A) Ordered according to their local phylogeny, 13 sequences of close (yellow) and distant (blue) homologs of CcNikZ-II and CjNikZ (pink) are aligned using MAFFT with coloring based on CLUSTALX at 25% conservation threshold. Three regions are shown: the variable loop (v-loop), the conserved arginine, and the conserved double histidine (HH)-prong with the numbers on the top indicating their relative positions. (B) Ordered according to their local phylogeny, sequences of selected members of the wild-type NiBP set are aligned and colored as before. Naming is based on their microorganism origins (cf. Table S3). The v-loop anatomy is shown (top). The v-loop section is between the left flank region and right flank region, which are strongly conserved. The v-loop section is comprised of the flexible core subsection (core v-loop) and the immobilized subsection (an α -helix). The core v-loop possesses a front and back to demark their different functions. The consensus sequence (bottom) is displayed, where '+' indicates the positions that do not have clear conservation.

sequences: P+L for the left flank region and LIF on the right flank region ('+' here indicating valine, isoleucine, and alanine) (Fig. 5B, Fig. S3). Sequence alignment with 10 NiBPs containing a double His prong revealed that the v-loop lengths varied from five to eight residues including the flexible core and rigid subsections. As shown in Fig. 5B, the last residue of the core v-loop and the first residue of the rigid subsection are aligned at the same (position 7) to reflect the varying boundary between the two subsections. Apart from CsNikZ, NgNikZ, CjNikZ, and CfNikZ that have an additional residue at position 0, the sequence EEH (positions 1–3) is the consensus sequence for the front of core v-loop. The majority of the CcNikZ-II homologs have either aspartate or glutamate at positions 1 and 2, followed by a histidine at position 3 (Fig. 5B). Since His26 from the CjNikZ v-loop was directly involved in Ni(II) coordination [13], we hypothesize that the presence of a carboxylate residue followed by a histidine in the core v-loop might confer higher Ni(II)-binding affinities.

Comparison of the core v-loops of 11 selected NiBPs (Fig. 5B) revealed that they differed both in length (5–8 residues) and composition. Because we were focused on understanding the role of the residues comprising CcNikZ-II's v-loop (TEDKYT), we decided to isolate for the effect of v-loop composition by keeping the following analysis restricted to a

6-residue window. We first analyzed the placement of the core v-loops relative to their Ni(II)-binding sites using structural models generated by ALPHAFOLD2 (Fig. 6, Figs S4 and S5). The v-loop regions of these proteins are shown in Fig. 6 with highlighted six-residue sequences selected for exchange with the CcNikZ-II core v-loop (TEDKYT). These selections were primarily based on considerations combining the critical positioning of a glutamate at the front of the core v-loop and structural confirmation that the residue positions are similar to the core v-loops of CcNikZ-II and CjNikZ (Fig. 1D). For example, BoNikZ from *Blautia obeum* possessed a glutamate (Glu22) with a histidine (His23) immediately downstream, so the remaining four residues were chosen to place Glu22 at the front of the core v-loop while preserving its second-from-the-start position: DEHGEI (Fig. 6). The structure shown in Fig. 6A helped to confirm this selection since the terminal isoleucine (Ile26) was positioned near the start of the rigid subsection, similar to the CcNikZ-II Thr64 (Fig. 1E). The CcNikZ-II engineered variant based on this selection was thus named *Bo* > *Cc* to denote the presence of DEHGEI in lieu of (>) TEDKYT. For VbNikZ from *Veillonellaceae bacterium* and TfNikZ from *Thermincola ferriacetica*, a glutamate was not present in the front of the core v-loop, so we simply used the first six residues following their left flank regions.

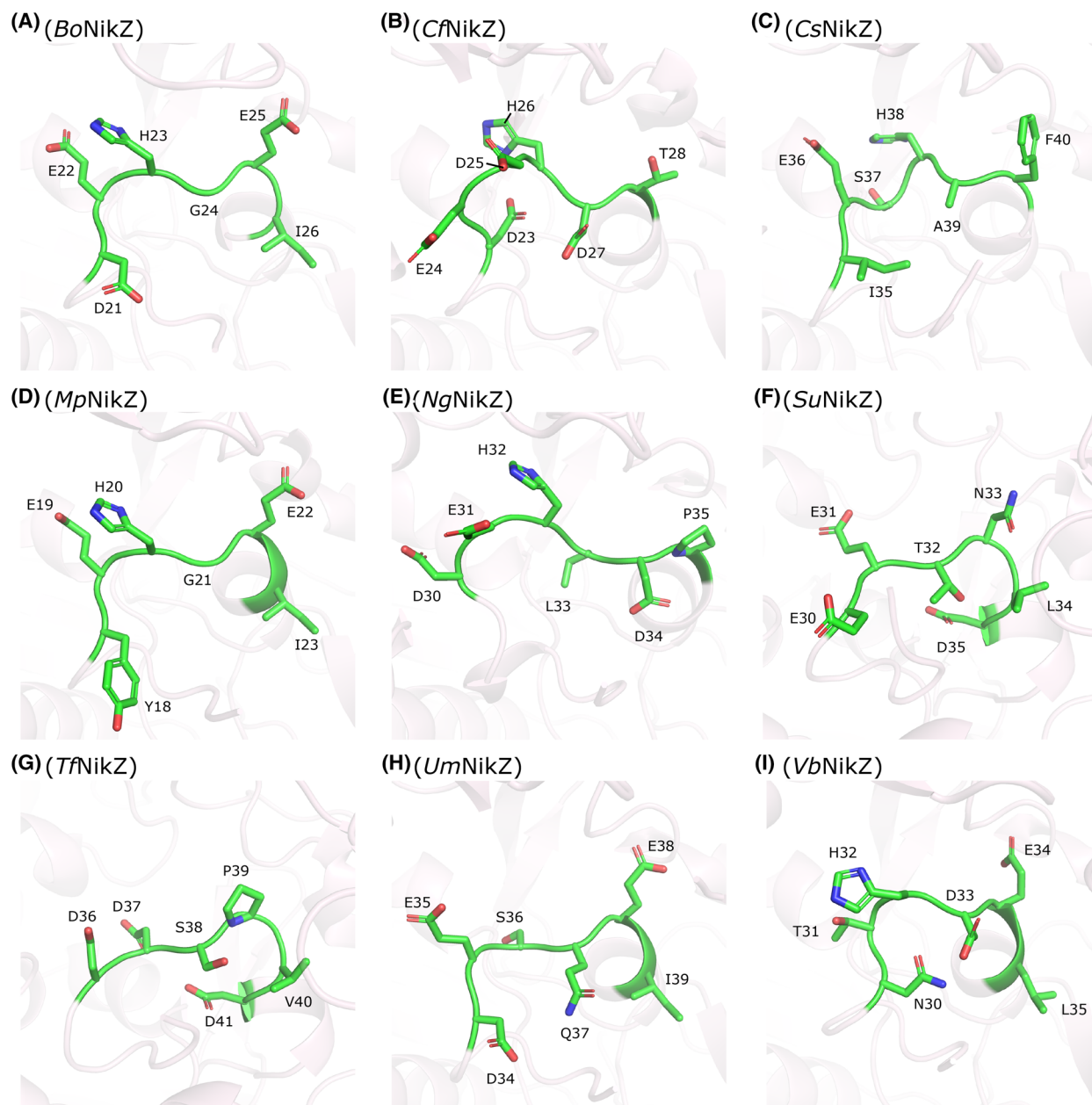


Fig. 6. Structural models of the CcNikZ-II homologs showing the core v-loops. Each panel displays the segment of the core v-loop (indicated by green carbons) that is used to substitute the CcNikZ-II core v-loop sequence TEDKTY. Residues are numbered according to their respective sequences. (A) *BoNikZ* (DEHGEI). (B) *CcNikZ* (DEDHDT). (C) *CsNikZ* (IESHAF). (D) *MpNikZ* (YEHGEI). (E) *NgNikZ* (DEHLDP). (F) *SuNikZ* (EETNLD). (G) *TcNikZ* (DDSPVD). (H) *UmNikZ* (DESQEI). (I) *VbNikZ* (NTHDEL).

Next, Ni(II)-binding affinity was determined for the purified wild-type NiBPs and CcNikZ-II variants with engineered v-loops (Fig. 7, Table S3, Figs S6–S8). Among the homologous NiBPs, CcNikZ-II showed the lowest Ni(II)-binding affinity (9.97 μM) (Fig. 7), which is likely due to the absence of a histidine residue and the presence of the hydrophilic Thr64 in the v-loop. The highest Ni(II)-binding affinities were detected

in *NgNikZ* from *Nocardiopsis gilva*, *BoNikZ*, and *MpNikZ* (from *Mogibacterium pumilum*) (K_D 0.06–0.22 μM) (Fig. 7, Table S3). *NgNikZ* was determined to have a sub-micromolar K_D value that should be carefully interpreted as an upper limit to its true K_D value since it was less than the concentration of the protein in the ITFQ assay [38]. This exceptionally high Ni(II)-binding affinity of *NgNikZ* is likely to be due to its

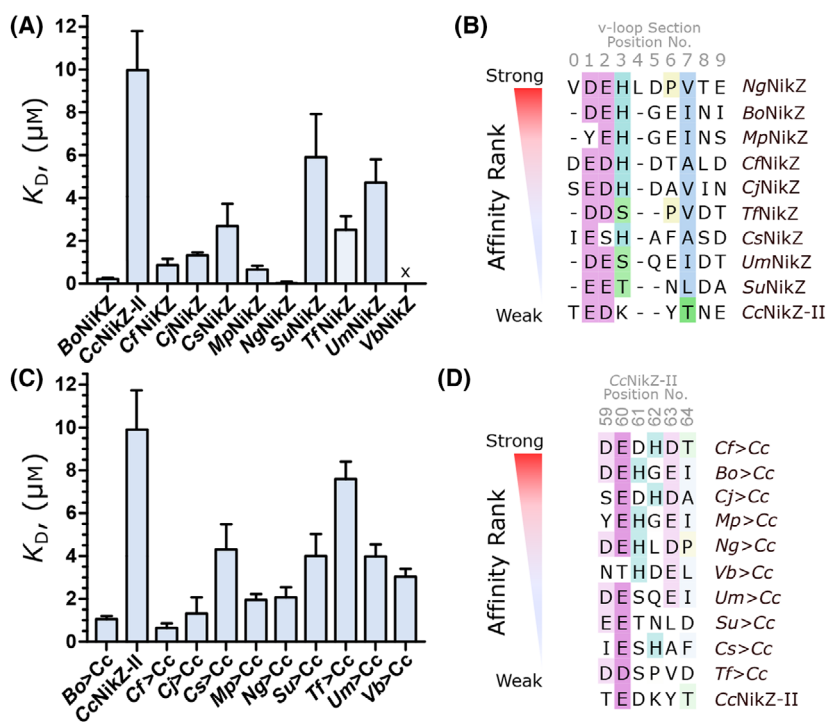


Fig. 7. Ni(II)-binding by purified wild type and engineered NiBPs. Ni(II)-binding affinities and ranking of purified wild type NiBPs (A, B); and engineered CcNikZ-II variants with different v-loops (C, D). The position numbering (gray) for (B) is the same as the numbering in Fig. 5B, and for (D) is based on the residue position in the CcNikZ-II sequence. Sequences are aligned based on MAFFT and colored according to CLUSTALX with a 25% conservation threshold. The position numbering of (D) is not the same as the numbering in Fig. 5B. 'x' denotes when a binding curve could not be generated (no observable change in fluorescence). Experimental triplicates were performed ($n=3$). Errors bars: SD.

origin in *Nocardiopsis gilva*, a gram-positive species discovered in hypersaline environments, which can grow in the presence of up to 18% NaCl (w/v) [39]. Furthermore, analysis of thermostability (T_m , melting points) of purified wild-type NiBPs revealed that TfNikZ from the thermophilic bacterium *Thermincola ferriacetica* (optimal growth at 60 °C) [40] had the highest T_m at 76 °C (Fig. S9). Together, these two proteins represent use cases for high salinity and high temperature metal recovery. Overall, ranking the NiBPs according to Ni(II) affinities (Fig. 7B) supported our hypothesis that high Ni(II)-binding affinity is associated with the presence of glutamate at position 2 followed by histidine at position 3. The sequences of the back of the core v-loops (positions 4–9) did not exhibit notable trends apart from position 7, where the presence of small hydrophobic residues (Val, Ile, Ala) correlated with higher Ni(II)-binding affinity.

Next, we engineered CcNikZ-II to prove that core v-loops have a degree of modularity whereby the replacement of the CcNikZ-II v-loop with sequences from higher-affinity homologs could improve its Ni(II)-binding affinity. As shown in Fig. 7C, all engineered variants of CcNikZ-II exhibited higher Ni(II)-binding affinity compared to wild-type CcNikZ-II. This is consistent with higher Ni(II)-binding affinities of the corresponding wild-type NiBPs. Ranking the engineered variants (Fig. 7D) further confirmed that the top five proteins for both wild type and engineered

NiBPs are based on the core v-loops from CfNikZ (from *Campylobacter fetus*), BoNikZ, CjNikZ, MpNikZ, and NgNikZ (tight binders). However, the order of ranking of the engineered CcNikZ-II variants was slightly different from that of the corresponding wild type NiBPs (Fig. 7), which may reflect the effect of the v-loop length and contribution from other residues present in their Ni(II)-binding sites that differ from CcNikZ-II (*vide infra*). We also noted that aspartates and glutamates at CcNikZ-II position 63 were also a common feature of the top five engineered CcNikZ-II variants (Fig. 7D). Since these residues are expected to be placed at the position of the Tyr63 side chain in the CcNikZ-II-binding site (Fig. 1D), it is plausible that aspartates and glutamates are better at contributing to a water network than tyrosine, thus improving Ni(II)-binding affinity of engineered variants (Fig. 1E). Overall, the largest improvements in Ni(II)-binding affinity were observed in the CcNikZ-II K62A and Cf>Cc (DEDHDT) variants indicating that v-loops do play an important role in Ni(II)-binding affinity of NiBPs.

Conclusions

Our approach to understanding the molecular mechanisms of Ni(II) binding and selectivity of NiBPs combined a bioinformatic analysis with biochemical and structural characterization of purified NiBPs. Structural analysis and Ala scanning mutagenesis of CcNikZ-II

revealed that Glu60 and His511 are required for high affinity Ni(II) binding suggesting an important role of the double His motif and v-loop residues in Ni(II) binding. Global phylogenetic analysis of NiBP sequences revealed the presence of at least 13 clusters of these proteins from which nine uncharacterized NiBPs with the double His prong and v-loops were selected for comparative analysis. The v-loops of these proteins were analyzed using ALPHAFOLD2 and showed the modest conservation of carboxylic residues (Glu and Asp) and His at the front of the six-residue core across several NiBPs, which correlated with high affinity Ni(II) binding by purified proteins (strong binders) without the requirement for chelators. Therefore, the C-terminal double His prong can be considered as the signature motif for the presence of chelator-free Ni(II)-binding activity in SBPs. Furthermore, all engineered CcNikZ-II variants with the v-loop core substituted with v-loops from strong binders showed increased Ni(II)-binding affinity, but the order of ranking did not match the ranking of wild type proteins, which is likely to be due to a combination of differences in the v-loop composition and secondary coordination sphere. As we have demonstrated in this study, NiBPs constitute promising protein engineering targets given their tuneable metal binding affinity. We note that this study only focused on the amino acid composition of the v-loop, but it is possible that changing v-loop length may also alter the NiBPs' metal-binding parameters. Additionally, another challenge to overcome will be controlling the effects of the second coordination sphere since the v-loop is only one component of the binding cavity, and there may be other residues that control the internal water network and placement of the binding residues in the first coordination sphere. While the literature describing this family of NiBPs has stagnated since the study of Lebrette *et al.*, we anticipate more interest in metalloprotein engineering towards novel bioprocesses to remediate and recover metals from aqueous mining and industrial streams due to rising interest in sustainable metal supply chains [13,24,41,42].

Materials and methods

Gene cloning and mutagenesis

The open-reading frames encoding for CcNikZ-II and other selected NiBPs (Tables S3 and S4) were synthesized (Twist Bioscience, San Francisco, CA, USA) without the start codon and without their native signal peptides, which were identified using SIGNALP 5.0 [43]. The open-reading frames (codon-optimized for expression in *E. coli*) were cloned into the expression vector p15TV-L (ID: 26093; AddGene,

Watertown, MA, USA) under the T7 promoter and in-frame with the N-terminal 6xHisTag (Twist Bioscience). Completed plasmids were transformed into the *E. coli* expression strain LOBSTR (#EC1001; Kerafast, Shirley, MA, USA) to minimize purification of endogenous Ni(II)-binding proteins, and then plated onto LB-agar (carbenicillin, 100 $\mu\text{g}\cdot\text{mL}^{-1}$). Colonies were inoculated into liquid LB, and the plasmids were miniprepmed (#PD300; GeneAid, New Taipei City, Taiwan) from overnight inoculants into MilliQ water and sequence verified at the ACGT Sequencing Facility (Toronto, ON, Canada). Glycerol stocks were prepared for storage at -80°C . All Ala substitution mutant proteins and engineered variants were generated through site-directed mutagenesis (Ranomics, Toronto, ON, Canada) with the TCG codon used for alanine substitution. Sequences were independently verified as described above.

Protein expression and purification

This protocol is based on prior work and was used to produce all proteins reported in this study [31]. Starter cultures were grown from glycerol stock in LB (carbenicillin, 100 $\mu\text{g}\cdot\text{mL}^{-1}$) for 16 h overnight at 37°C with shaking. Expression cultures were started by prewarming TB media (carbenicillin, 100 $\mu\text{g}\cdot\text{mL}^{-1}$) to 37°C before 5% v/v inoculation with the starter culture, then grown for 6 h at 37°C , then addition of IPTG (#IPT002; BioShop) to 0.4 mM for induction. The expression cultures were then transferred to 16°C and grown for 16 h overnight with shaking, then pelleted with centrifugation and transferred to conical vials for one overnight freeze-thaw cycle at -20°C . Frozen cell pellets were thawed and resuspended in binding buffer (10 mM HEPES, 500 mM NaCl, 5 mM imidazole, pH 7.2) to a final volume of 50–100 mL, followed by addition of 0.25 g lysozyme (#LYS702; BioShop, Burlington, ON, Canada). Cell pellet mixtures were sonicated for 15 min (Q700 Sonicator; Qsonica, Newtown, CT, USA) and clarified by centrifugation. The soluble layer (supernatant) was applied to a cobalt-charged (Co-)NTA resin (#88221; Thermo Fisher) pre-equilibrated with binding buffer in a gravity-column set-up. Bound proteins were cleansed with wash buffer (10 mM HEPES, 500 mM NaCl, 25 mM imidazole, pH 7.2) and collected with elution buffer (10 mM HEPES, 500 mM NaCl, 250 mM imidazole, pH 7.2). Protein concentrations were determined by Bradford assay, and protein purity was determined by SDS/PAGE analysis (Fig. S1) and densitometry on IMAGE LAB 6.0 (Bio-Rad).

Eluted protein was combined with in-house purified TEV protease (100:1 protein-to-TEV) an DTT to 1 mM (#TCE101; BioShop), then transferred to a 10 kDa MWCO dialysis bag (#68100; Thermo Fisher) for dialysis in 4 L dialysis buffer (10 mM HEPES, 1 mM DTT, 1 $\text{g}\cdot\text{L}^{-1}$ Chelex 100, pH 7.2) at 4°C with gentle stirring for 24 h. Dialyzed samples were then applied to a cobalt-charged NTA resin twice and transferred to a 10 kDa MWCO dialysis bag for

dialysis in 4 L activity buffer (AB, 10 mM HEPES, pH 7.2) at 4 °C with gentle stirring for 24 h and transferred to fresh AB for another 24 h. Purity was re-confirmed by SDS/PAGE analysis. Finally, they were flash-frozen by liquid nitrogen for storage at −80 °C.

Protein crystallization and x-ray crystallography

The apo CcNikZ-II crystal was grown at room temperature using the vapor diffusion sitting drop method solution containing 30 mg·mL^{−1} protein and the reservoir solution 0.1 M HEPES buffer (pH 7.5), 30% (w/v) polyethylene glycol 1000. The crystal was cryoprotected using paratone oil. Diffraction data at 100 K were collected at a home source Rigaku Micromax 007-HF/R-Axis IV system and data was processed using HKL3000 [44]. The structure was solved by molecular replacement, using the structure of apo CjNikZ (PDB: 4OET) [13] and the ccr4 Mr Bump program [45]. Model building and refinement were performed using PHENIX.REFINE [46] and COOT [47]. Geometry was validated using the wwPDB validation server. Atomic coordinates were deposited in the PDB with accession code 8EFZ.

Phylogenetic and structural analyses

Under the broad family IPR039424 (solute binding protein family 5) in the InterPro database, amino acid sequences from family IPR030678 (peptide/Ni(II)-binding protein, MppA-type) were extracted into a FASTA file on April 27, 2019, which contained approximately 150 000 members [36]. Sequence redundancy was reduced through the CD-HIT algorithm using an identity cutoff of 0.6 (60%) and default settings [35]. Amino acid sequences of EcNikA, SaCntA, YpYntA, SaNikA, VpNikA, HhNikA, CjNikZ, and CcNikZ-II were then manually added (if not already present) to the newly reduced FASTA file, which then had 4561 members. An alignment was performed using the MAFFT v7 algorithm through GENEIOUS v8.1.9 with default settings, and a phylogenetic tree was then built using FASTTREE v2.1.11 with the Gamma 20 optimization setting [48]. Site 1 and Site 2 were extracted from the above tree and imported into JALVIEW v.2.11.2.2 for closer analysis. Sequences of interest were further extracted and re-aligned again with MAFFT v7. A phylogenetic tree was built based on and default settings to examine binding motif conservation.

PYMOL v.2.5.2. was used to visualize protein structures. CcNikZ-II's structure was predicted on SWISS-MODEL using the apo CjNikZ structure (PDB: 4OET) as the homology template [49]. It was also predicted using ALPHAFOLD2 through the COLABFOLD server with the amber settings and without a template [32,33]. Structural superimposition was performed using built-in PYMOL functions and rmsd values (Å) were automatically computed. These ALPHAFOLD2 prediction and PYMOL procedures used for CcNikZ-II were applied to all members of the engineered variant set to predict and

compare their apo structures with apo CcNikZ-II's structure reported in this study. Electrostatic surface representations were generated using the APBS biomolecular solvation software suite [50].

Metal-affinity determination by microITFQ-LTA

The microplate-based intrinsic tryptophan fluorescence quenching (or 'ITFQ assays' for brevity) is based on prior work and was used to characterize all proteins reported in this study [31]. ITFQ assays were performed in black, opaque 96-well microplates (#655076; Greiner Bio-One, Monroe, NC, USA) using the Infinite® M200 (Tecan) plate reader with the settings: fluorescence top-read, 25 °C, λ_{ex} = 280 nm, λ_{em} = 380 nm, and a manual gain of 100. The general procedure used to obtain binding curves by direct titration first required 150 μL of 0.4 μM CcNikZ-II in AB to be added to the appropriate number of wells and equilibrated for 20 min with shaking (orbital, 3 mm) to 25 °C. The baseline fluorescence was monitored to ensure equilibration, from which one endpoint reading was made as the blank (F₀). Titrations of metal working solution were incrementally added. Specifically, at each titration increment, a small aliquot (1–20 μL) was added by multi-channel pipette, then equilibrated for 3 min with shaking (orbital, 3 mm). For larger aliquots (10–20 μL), the samples were equilibrated for at least 5 min with shaking (orbital, 3 mm), or until the fluorescence stabilized. At each titration step, after the 3–5 min equilibration period, the fluorescence was measured (F). This titration procedure was repeated until saturation of the fluorescence signal was observed.

Mixing by pipette was avoided to prevent the introduction of air bubbles or the unintentional removal of protein due to droplets sticking to the inside of pipette tips. To account for any dilution effects, we included a triplicate negative control where the protein was titrated with AB (no added metal) using the same volumes added at each titration step. The inner-filter effect was not observed for the concentrations of metal working solutions used.

$$\frac{F_{\text{obs}}}{F_{\text{max}}} = \frac{[P]_{\text{tot}} + [M]_{\text{tot}} + K_D - \sqrt{([P]_{\text{tot}} + [M]_{\text{tot}} + K_D)^2 - 4[P]_{\text{tot}}[M]_{\text{tot}}}}{2[P]_{\text{tot}}} \quad (1)$$

$$\sigma_x = \sqrt{\sigma_a^2 + \sigma_b^2} \quad (2)$$

The apparent dissociation constant (K_D) and the maximum fluorescence change (F_{max}) were determined by fitting a 1 : 1 stoichiometry mass-action kinetic model (Eqn 1) to the data (Table S2) [51]. Here, [M]_{tot} and [P]_{tot} represent the total metal and protein concentration in the well at each

titration step. The observed fluorescence change F_{obs} ($F_0 - F$) is calculated by subtracting the fluorescence of the sample when metal is added (F) at each titration step from the fluorescence when no metal is added (F_0). Experiments were performed in triplicates. The standard deviation of the resulting F_{obs} is therefore calculated by error propagation (Eqn 2). The standard deviations of K_D and F_{max} are computed by the `scipy.optimize` function in PYTHON where this data was analyzed [52]. Plots were generated using `matplotlib` in PYTHON [53].

Acknowledgements

The authors thank (a) Dr Anna Khusnutdinova and Dr Sofia Lemak for advice and support, and (b) Howie Lam for preliminary protein solubility trials. This work was supported by the Ontario Ministry of Economic Development, Job Creation and Trade through the Elements of Bio-mining ORF-RE program. PD is grateful to be a recipient of Ontario Graduate Scholarships. This work was also in part performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344 (LLNL-JRNL-861367).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

PD contributed to the conceptualization, investigation, methodology, formal analysis, visualization, writing—original draft, and review and editing. PJS contributed to the investigation, formal analysis, and writing—original draft. EE contributed to the investigation. AS contributed to the supervision and funding acquisition. RM contributed to the review and editing, supervision, project administration, and funding acquisition. AFY contributed to the writing—original draft, review and editing, supervision, project administration, and funding acquisition.

Peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.17125>.

Data availability statement

The atomic coordinates and structure factors of the model described in this manuscript were deposited

into the Protein Data Bank with the accession code [8EFZ](#).

References

- Waldron KJ, Rutherford JC, Ford D & Robinson NJ (2009) Metalloproteins and metal sensing. *Nature* **460**, 823–830.
- Foster AW, Young TR, Chivers PT & Robinson NJ (2022) Protein metalation in biology. *Curr Opin Chem Biol* **66**, 102095.
- Begg SL (2019) The role of metal ions in the virulence and viability of bacterial pathogens. *Biochem Soc Trans* **47**, 77–87.
- Chandrangsu P, Rensing C & Helmann JD (2017) Metal homeostasis and resistance in bacteria. *Nat Rev Microbiol* **15**, 338–350.
- Haferburg G & Kothe E (2007) Microbes and metals: interactions in the environment. *J Basic Microbiol* **47**, 453–467.
- Chivers PT (2015) Nickel recognition by bacterial importer proteins. *Metallomics* **7**, 590–595.
- Zhang P (2013) Structure and mechanism of energy-coupling factor transporters. *Trends Microbiol* **21**, 652–659.
- Rees DC, Johnson E & Lewinson O (2009) ABC transporters: the power to change. *Nat Rev Mol Cell Biol* **10**, 218–227.
- Scheepers GH, Lycklama a Nijeholt JA & Poolman B (2016) An updated structural classification of substrate-binding proteins. *FEBS Lett* **590**, 4393–4401.
- Berntsson RP-A, Smits SHJ, Schmitt L, Slotboom D-J & Poolman B (2010) A structural classification of substrate-binding proteins. *FEBS Lett* **584**, 2606–2617.
- Saier MH Jr, Tran CV & Barabote RD (2006) TCDB: the transporter classification database for membrane transport protein analyses and information. *Nucleic Acids Res* **34**, D181–D186.
- Tam R & Saier MH (1993) Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol Rev* **57**, 320–346.
- Lebrette H, Brochier-armanet C, Zambelli B, de Reuse H, Borezée-Durant E, Ciurli S & Cavazza C (2014) Promiscuous nickel import in human pathogens: structure, thermodynamics, and evolution of extracytoplasmic nickel-binding proteins. *Structure* **22**, 1421–1432.
- Chandravanshi M, Tripathi SK & Kanaujia SP (2021) An updated classification and mechanistic insights into ligand binding of the substrate-binding proteins. *FEBS Lett* **595**, 2395–2409.
- Howlett RM, Hughes BM, Hitchcock A & Kelly DJ (2012) Hydrogenase activity in the foodborne pathogen *Campylobacter jejuni* depends upon a novel ABC-type nickel transporter (NikZYXWV) and is SlyD-independent. *Microbiology* **158**, 1645–1655.

- 16 Lebrette H, Borezée-Durant E, Martin L, Richaud P, Boeri Erba E & Cavazza C (2015) Novel insights into nickel import in *Staphylococcus aureus*: the positive role of free histidine and structural characterization of a new thiazolidine-type nickel chelator†. *Metallomics* **7**, 613–621.
- 17 Benoit SL, Seshadri S, Lamichhane-Khadka R & Maier RJ (2013) *Helicobacter hepaticus* NikR controls urease and hydrogenase activities via the NikABDE and HH0418 putative nickel import proteins. *Microbiology* **159**, 136–146.
- 18 Shaik MM, Cendron L, Salamina M, Ruzzene M & Zanotti G (2014) *Helicobacter pylori* periplasmic receptor CeuE (HP1561) modulates its nickel affinity via organic metallophores. *Mol Microbiol* **91**, 724–735.
- 19 Park K-S, Iida T, Yamaichi Y, Oyagi T, Yamamoto K & Honda T (2000) Genetic characterization of DNA region containing the trh and ure genes of *Vibrio parahaemolyticus*. *Infect Immun* **68**, 5742–5748.
- 20 Hughes A, Wilson S, Dodson EJ, Turkenburg JP & Wilkinson AJ (2019) Crystal structure of the putative peptide-binding protein AppA from *Clostridium difficile*. *Acta Crystallogr F* **75**, 246–253.
- 21 Wegner SV, Boyaci H, Chen H, Jensen MP & He C (2009) Engineering a uranyl-specific binding protein from NikR. *Angew Chem Int Ed* **48**, 2339–2341.
- 22 Jia X, Ma Y, Bu R, Zhao T & Wu K (2020) Directed evolution of a transcription factor PbrR to improve lead selectivity and reduce zinc interference through dual selection. *AMB Express* **10**, 67.
- 23 Zhou L, Bosscher M, Zhang C, Özçubukçu S, Zhang L, Zhang W, Li CJ, Liu J, Jensen MP, Lai L *et al.* (2014) A protein engineered to bind uranyl selectively and with femtomolar affinity. *Nat Chem* **6**, 236–241.
- 24 Dong Z, Mattocks JA, Deblonde GJ-P, Hu D, Jiao Y, Cotruvo JA & Park DM (2021) Bridging hydrometallurgy and biochemistry: a protein-based process for recovery and separation of rare earth elements. *ACS Cent Sci* **7**, 1798–1808.
- 25 Keshav V, Franklyn P & Kondiah K (2019) Recombinant fusion protein PbrD cross-linked to calcium alginate nanoparticles for Pb remediation. *ACS Omega* **4**, 16816–16825.
- 26 Choi TS & Tezcan FA (2022) Overcoming universal restrictions on metal selectivity by protein design. *Nature* **603**, 522–527.
- 27 Irving H & Williams RJP (1953) The stability of transition-metal complexes. *J Chem Soc* 3192–3210.
- 28 Foster AW, Osman D & Robinson NJ (2014) Metal preferences and metallation. *J Biol Chem* **289**, 28095–28103.
- 29 Dudev T & Lim C (2014) Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. *Chem Rev* **114**, 538–556.
- 30 Waldron KJ & Robinson NJ (2009) How do bacterial cells ensure that metalloproteins get the correct metal? *Nat Rev Microbiol* **7**, 25–35.
- 31 Diep P, Mahadevan R & Yakunin AF (2020) A microplate screen to estimate metal-binding affinities of metalloproteins. *Anal Biochem* **609**, 113836.
- 32 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
- 33 Mirdita M, Schütze K, Moriawaki Y, Heo L, Ovchinnikov S & Steinegger M (2022) ColabFold: making protein folding accessible to all. *Nat Methods* **19**, 679–682.
- 34 Holm L & Rosenström P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.
- 35 Li W & Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659.
- 36 Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S *et al.* (2021) The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* **49**, D344–D354.
- 37 Price MN, Dehal PS & Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490.
- 38 Young TR & Xiao Z (2021) Principles and practice of determining metal–protein affinities. *Biochem J* **478**, 1085–1116.
- 39 Li W-J, Kroppenstedt RM, Wang D, Tang S-K, Lee J-C, Park D-J, Kim C-J, Xu L-H & Jiang C-L (2006) Five novel species of the genus *Nocardiopsis* isolated from hypersaline soils and emended description of *Nocardiopsis salina* Li *et al.* 2004. *Int J Syst Evol Microbiol* **56**, 1089–1096.
- 40 Zavarzina DG, Sokolova TG, Tourova TP, Chernyh NA, Kostrikin NA & Bonch-Osmolovskaya EA (2007) *Thermincola ferriacetica* sp. nov., a new anaerobic, thermophilic, facultatively chemolithoautotrophic bacterium capable of dissimilatory Fe(III) reduction. *Extremophiles* **11**, 1–7.
- 41 Capeness MJ & Horsfall LE (2020) Synthetic biology approaches towards the recycling of metals from the environment. *Biochem Soc Trans* **48**, 1367–1378.
- 42 Urbina J, Patil A, Fujishima K, Paulino-Lima IG, Saltikov C & Rothschild LJ (2019) A new approach to biomining: bioengineering surfaces for metal recovery from aqueous solutions. *Sci Rep* **9**, 16422.
- 43 Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G & Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420–423.

- 44 Minor W, Cymborowski M, Otwinowski Z & Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes. *Acta Crystallogr D* **62**, 859–866.
- 45 Keegan RM, McNicholas SJ, Thomas JMH, Simpkin AJ, Simkovic F, Uski V, Ballard CC, Winn MD, Wilson KS & Rigden DJ (2018) Recent developments in MrBUMP: better search-model preparation, graphical interaction with search models, and solution improvement and assessment. *Acta Crystallogr D* **74**, 167–182.
- 46 Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung L-W, Jain S, McCoy AJ *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D* **75**, 861–877.
- 47 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of coot. *Acta Crystallogr D* **66**, 486–501.
- 48 Katoh K, Misawa K, Kuma K & Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059–3066.
- 49 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296–W303.
- 50 Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K *et al.* (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112–128.
- 51 van de Weert M & Stella L (2011) Fluorescence quenching and ligand binding: a critical discussion of a popular methodology. *J Mol Struct* **998**, 144–150.
- 52 Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J *et al.* (2020) SciPy 1.0:

fundamental algorithms for scientific computing in python. *Nat Methods* **17**, 261–272.

- 53 Hunter JD (2007) Matplotlib: a 2D graphics environment. *Comput Sci Eng* **9**, 90–95.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. SDS/PAGE analysis of purified NiBP proteins used in this study.

Fig. S2. Nickel (Ni^{2+}) binding curves of the alanine mutants.

Fig. S3. Full multiple sequence alignment of wild-type NiBPs.

Fig. S4. V-loop structural alignment of ALPHAFOLD2-predicted structures.

Fig. S5. Structural alignment of v-loops from the crystal structure of CcNikZ-II and structural models generated using SWISS-MODEL (A) or ALPHAFOLD2 (B).

Fig. S6. Ni^{2+} -binding curves of the wild-type NiBPs characterized in this study.

Fig. S7. Ni^{2+} -binding curves of the engineered CcNikZ-II variants.

Fig. S8. Ni^{2+} -binding curves of purified wild-type NiBPs and engineered CcNikZ-II proteins with the lowest K_D values (highest affinity) used in this study.

Fig. S9. Melting points of NiBPs.

Table S1. X-ray crystallographic statistics for the structure of CcNikZ-II.

Table S2. Nickel-binding parameters of the wild type and engineered NiBPs.

Table S3. NiBPs characterized in this study.

Table S4. Amino acid sequences of CcNikZ-II and other NiBPs characterized in this study.