Structural insights into inherited anemia CDA-I: disease-associated mutations disrupt CDIN1-Codanin1 complex

Martin Stojaspal[1,2], Tomáš Brom[1], Ivona Nečasová[1,2], Tomáš Janovič[1,†], Pavel Veverka[1,2], Lukáš Uhrík[3], Lenka Hernychova[3], and Ctirad Hofr[1*]

[1]LifeB, FGP – NCBR, Faculty of Science, Masaryk University, Kamenice 753/5, Brno 625 00, Czech Republic

[2]Institute of Biophysics of the Czech Academy of Sciences, Scientific Incubator, Kralovopolska 135, Brno 612 00, Czech Republic

[3]Research Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Zluty kopec 7, Brno 656 53, Czech Republic

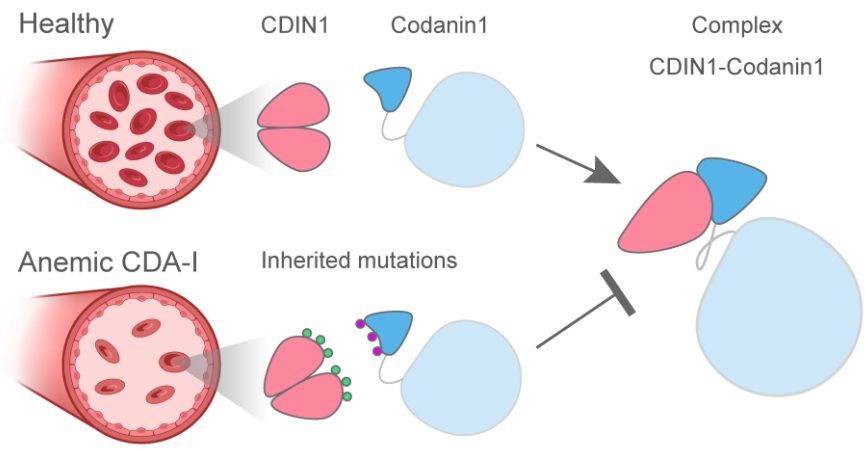
[†]Present address: Institute for Quantitative Health Science and Engineering, Michigan State University, 775 Woodlot Dr, East Lansing, MI 48824, USA

[*]Correspondence: Tel: +420 549 495 952, Fax: +420 549 492 640, E-mail: hofr@sci.muni.cz

Keywords

Rare diseases, congenital dyserythropoietic anemia, Codanin1, CDIN1, C15orf41, protein-protein interaction, structure predictions

Graphical Abstract



Highlights

- Full-length CDIN1 preferentially forms dimers, Codanin1$_{Cterm}$ monomers
- CDIN1 binds Codanin1$_{Cterm}$ in equimolar ratio with nanomolar affinity
- Determined interacting regions of CDIN1 and Codanin1$_{Cterm}$ contain mutations associated with CDA-I disease
- CDA-I-related mutations disrupt the binding of CDIN1 and Codanin1$_{Cterm}$

Abstract

Congenital dyserythropoietic anemia type I (CDA-I) is a rare hereditary disease characterized by ineffective erythropoiesis and associated mutations in two proteins – Codanin1 and CDIN1. The primary role of Codanin1 is nucleosome assembly regulation through interaction with ASF1. The role of recently discovered CDIN1 remains unknown, but CDIN1 has been known to interact directly with the C-terminus of Codanin1. Despite the critical role of identified mutations in Codanin1 and CDIN1, the effects of CDA-I-related mutations at the molecular level have not been elucidated.

Here, we reconstruct the structural envelopes of CDIN1 and Codanin1, determine stoichiometry, define essential interacting regions, and quantify mutual affinity. We demonstrate that the anemia-associated mutations disturb CDIN1 and Codanin1 binding. Our findings present new insights into the structure of Codanin1 and CDIN1 and the functional effects of disease-associated mutations as the next step in unraveling the molecular etiology of CDA-I disease.

Introduction

Congenital dyserythropoietic anemia (CDA) is a rare hereditary disease manifesting by abnormal erythroblast morphology and ineffective erythropoiesis in the bone marrow[1]. CDA type I (CDA-I) incidence is 1 per ~207 000 live births according to recent allele frequency analysis[2]. CDA-I causes spleen enlargement and iron overload. At the cytological level, CDA-I presents increased mean red blood cell volume in 75% of all erythroblasts. Additionally, CDA-I induces binucleate erythroblasts (up to 7%) and internuclear chromatin bridges (up to 3%). Most strikingly, electron microscopy revealed spongy heterochromatin with a "Swiss-cheese" like appearance in up to 60% of cells[3]. Moreover, cell cycle analyses of patient samples showed an accumulation of erythroblasts in S-phase, which suggests replication problems.

From a genetic point of view, CDA-I is an autosomal recessive disease caused by biallelic mutations in genes *CDAN1* and *CDIN1* (originally called *C15orf41*), which account for approximately 90% of all CDA-I cases[4]. *CDAN1* was first identified as the gene causing CDA-I disease in a group of Israeli Bedouins[5]. *CDAN1* is essential for cell survival as mice embryos with an artificially disrupted gene perished in the early stages of development[6]. The highly conserved *CDAN1* gene[5] encodes Codanin1, a 134 kDa protein that negatively regulates the shuttling of newly synthesized histones from the cytoplasm to the nucleus by histone chaperone Asf1[7].

The second causative gene – *CDIN1,* was identified by whole-genome sequencing of individuals from CDA-I disease-affected pedigrees[8]. *CDIN1* encodes the 32 kDa CDAN1-interacting nuclease 1 (CDIN1) of unknown function. Although CDIN1 was originally predicted as a restriction nuclease[8], no evidence of nuclease activity has yet been published. Gene expression arrays revealed that *CDIN1* is most extensively transcribed in hematopoietic stem cells, B lymphoblasts, cardiomyocytes, and fetal liver, suggesting the essentiality of CDIN1 in hematopoiesis[9].

Co-immunoprecipitation and immunofluorescence studies demonstrated that the two crucial anemia-associated proteins CDIN1 and Codanin1 bind directly[2,4,10]. Additionally, Swickley et al. and Shroff et al. independently found that CDIN1 binds the C-terminal part of Codanin1 ($Codanin1_{Cterm}$) comprising amino acids 1001-1227[4,10]. The most recent genetic testing and molecular diagnoses revealed eight missense mutations in $Codanin1_{Cterm}$ and six mutated amino acids in CDIN1 related to CDA-I progression[2,11], but it remains uncertain whether CDIN1 contributes to histone trafficking and how CDIN1 affects other Codanin1 functions and CDA-I pathology.

Despite the known essential role of CDIN1 and $Codanin1_{Cterm}$ in CDA-I pathology, their interacting regions have not been mapped and defined. Furthermore, pathological CDA-I mutations have not been linked to CDIN1-Codanin1 interaction. Here we report the interacting regions critical for the binding of CDIN1 and Codanin1. We present SAXS structural envelopes of CDIN1 and $Codanin1_{Cterm}$. Moreover, our quantitative studies determined that CDIN1 and $Codanin1_{Cterm}$ bind with a high affinity. Finally, we showed that the CDA-I-associated mutations in interacting regions disturb CDIN1-Codanin1

complex. The revealed structure-function relationship of disease-related mutations is essential for determining the molecular mechanism of CDIN1 and Codanin1 and for the further development of biological treatment of CDA-I.

Results

CDIN1 preferentially occurs as a homodimer and Codanin1$_{Cterm}$ as a monomer

To characterize individual proteins and describe their interaction, we produced codon-optimized mammalian proteins CDIN1 and Codanin1$_{Cterm}$ using a bacterial system. We purified CDIN1 and Codanin1$_{Cterm}$ using a three-step purification with a final size-exclusion chromatography (SEC) step. We used SDS-PAGE and MALDI-TOF MS to analyze protein purity and identity (Fig. S1). Subsequent spectrometric and spectroscopic measurements confirmed that we prepared pure proteins that were adequately folded, homogenous, and stable (Fig. S2).

To determine the most probable secondary structure composition and spatial arrangements of individual proteins, we used AlphaFold predicted models combined with advanced biophysical methods. We examined predicted structures of CDIN1 (Q9Y2V0) and Codanin1 (Q8IWY9) available in the AlphaFold Protein Structure Database. The CDIN1 structural model was predicted with high confidence – the per-residue confidence score (pLDDT) was higher than 90 for 257 of the total 281 amino acids. CDIN1 preferentially forms α-helices and sporadically short β-sheets (Fig 1A). The 59% α-helix and 13% β-sheet content in the CDIN1 prediction model closely match experimentally obtained values of 45% α-helix and 13% β-sheet calculated from circular dichroism (CD) measurements (Fig. S2A, B). Similarly, Codanin1$_{Cterm}$ structure was calculated with high confidence with pLDDT higher than 80 for 157 of the total 222 amino acids. Codanin1$_{Cterm}$ consists of α-helices forming a pair of four-helix bundles according to the predicted structural model (Fig. 1B). The predicted 53% α-helix content for Codanin1$_{Cterm}$ is in excellent agreement with the 55% content of helical structures calculated from CD measurements (Fig. S2A, C). Altogether, structural predictions and our experimental CD data suggest that both CDIN1 and Codanin1$_{Cterm}$ predominantly arrange into alpha helices.

To investigate the oligomeric states of individual proteins, we performed native gel electrophoresis on purified proteins. CDIN1 and Codanin1$_{Cterm}$ were mixed in reducing or non-reducing loading dye and separated in a native gel to visualize possible multimers. Under non-reducing conditions preserving disulfide bonds, we observed four distinct bands in both proteins (Fig. S2F). CDIN1 and Codanin1$_{Cterm}$ bands in non-reducing conditions were significantly shifted to higher molecular weights compared to those observed in reducing conditions (Fig. S2F), suggesting the presence of several oligomeric forms for both proteins.

We then quantified the distribution of CDIN1 and Codanin1$_{Cterm}$ oligomeric states by combining size-exclusion chromatography and multiangle light scattering (SEC-MALS) to separate complexes by their size and shape and to obtain the molecular weight of each form. The strongest CDIN1 signal corresponds to a molecular weight of 69 kDa, which matches the expected weight of a homodimer (Fig. 1C). We also detected a signal at 137 kDa that likely indicates a homotetramer. The broad peak with a maximum of 49 kDa might reflect the equilibrium between monomeric and dimeric CDIN1. Similarly, the peak at 115 kDa might describe a transition state between dimeric and tetrameric CDIN1. In contrast, the most prominent Codanin1$_{Cterm}$ signal corresponds to a molecular weight of 26 kDa (Fig. 1D), which matches the monomeric state. Two marginal peaks corresponding to comparable molecular weights 50 and 55 kDa suggest the presence of dimeric forms with different conformations of Codanin1$_{Cterm}$. In conclusion, our SEC-MALS measurements and electrophoretic assay suggest that CDIN1 adopts predominantly homodimeric arrangements in solution, whereas Codanin1$_{Cterm}$ occurs primarily as a monomer.

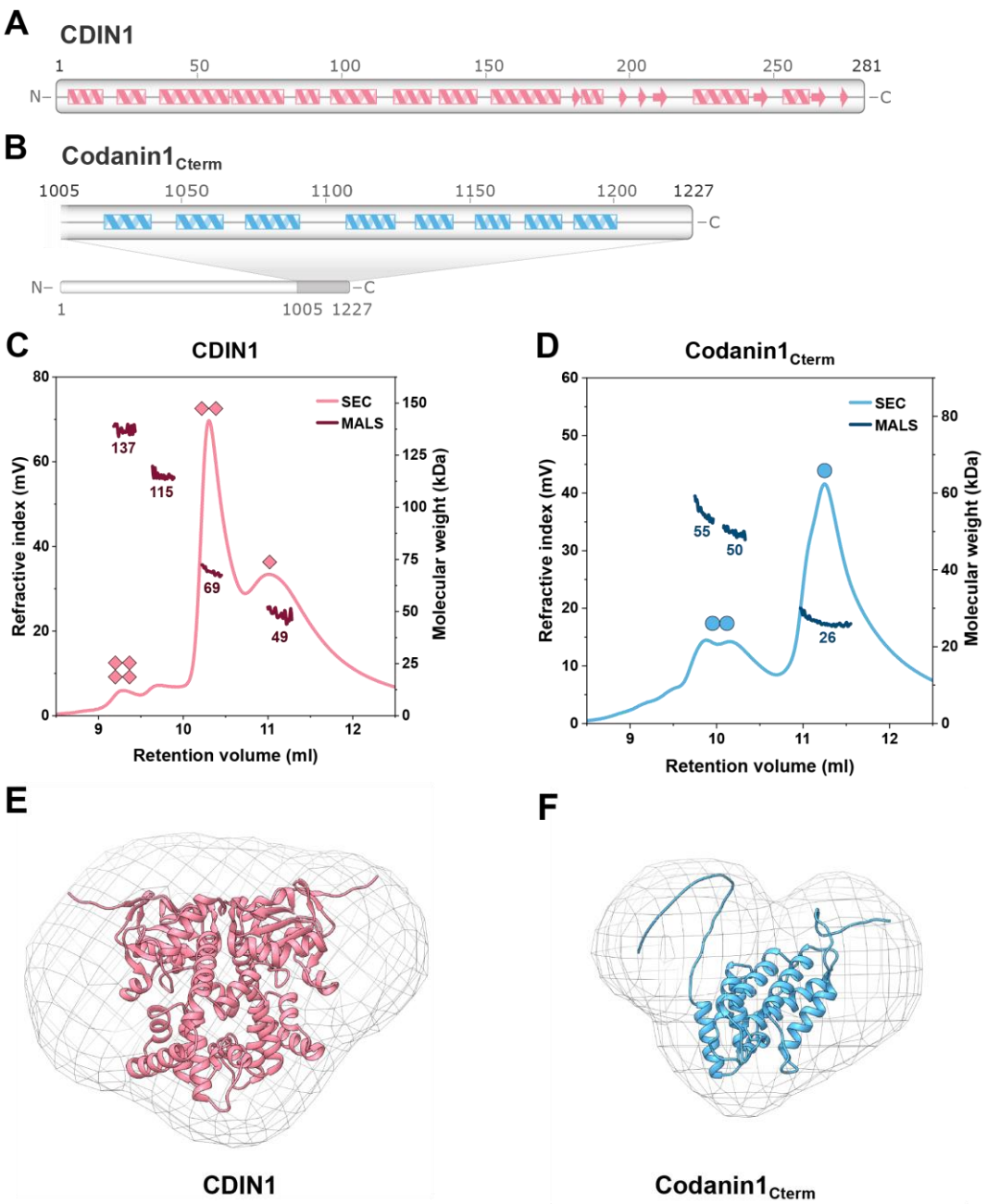


Figure 1. Human CDIN1 preferentially forms a dimer and Codanin1$_{Cterm}$ monomer.
(A, B) Prevalently α-helical secondary structure of CDIN1 and Codanin1$_{Cterm}$ predicted by AlphaFold. **(C, D)** SEC-MALS shows the dominant dimeric arrangement of CDIN1 and mainly monomeric occurrence of Codanin1$_{Cterm}$. **(E, F)** SAXS structural envelopes accommodate CDIN1 dimer and Codanin1$_{Cterm}$ monomer.

To gain insight into the overall structure of CDIN1 and Codanin1$_{Cterm}$ in solution, we carried out measurements by small-angle X-ray scattering (SAXS) after SEC separation. In agreement with the SEC-MALS data, the structural parameters determined by scattering measurements in solution show that CDIN1 is dimeric and Codanin1$_{Cterm}$ monomeric (Table S1). The Guinier plots for both proteins are linear (Fig. S3E, S4E), indicating the absence of aggregation or inter-particle interference in the solution. The calculated radius of gyration (R_g) from both the Guinier approximation and from the concentration dependence of distance distribution function P(r) are identical within the precision of the measurement (Table S1). Furthermore, the Kratky plot for CDIN1 is a bell-shaped curve, indicating a folded protein (Fig. S3F). Similarly, the Kratky plot for Codanin1$_{Cterm}$ shows a bell-shaped

tendency influenced by a higher ratio of unstructured parts in agreement with AlphaFold predictions, demonstrating a folded core with flanking disordered regions (Fig. S4F).

For CDIN1, we observed a single broad peak in the SEC elution profile. When we analyzed SAXS data recorded at the SEC peak, we obtained a mean R_g value of 2.95 ± 0.06 nm (Fig. S3A, Table S1). Additionally, we used experimental SAXS data to construct the three-dimensional shape of CDIN1 using the ATSAS software package[12]. We reconstructed a low-resolution molecular envelope by averaging ten individual ab initio bead models. When we superimposed the calculated structural envelope onto the predicted dimeric structure of CDIN1, we observed that the dimer embedded well into the envelope (Fig. 1E). To compare theoretical scattering profiles generated from the predicted dimer structure with our experimental SAXS data, we employed CRYSOL[13]. The fit between the calculated scattering pattern and the experimental scattering curve (χ^2 = 2.1) demonstrates the high relevance of the calculated data. In contrast, when we tried to fit monomeric CDIN1 into the SAXS envelope, the overall fit was poor (χ^2 = 8.1) (Fig. S3D).

For Codanin1$_{Cterm}$, we recorded the SEC elution profile revealing two well-separated fractions (Fig. S4A-B, S5A-B). For the particles eluted in peak fraction with higher retention volumes, we calculated a R_g of 2.66 ± 0.07 nm from SAXS data using Primus software (Table S1). Subsequently, comparing the experimental scattering curve and scattering pattern calculated from the spatial model of Codanin1$_{Cterm}$, we observed a perfect agreement (χ^2 = 1.2) (Fig. S4D). Finally, we superimposed the heart-shaped SAXS envelope of Codanin1$_{Cterm}$ and the predicted monomer structure. We observed that the monomer fits into the SAXS envelope (Fig. 1F). The monomeric fraction prevailed even at high Codanin1$_{Cterm}$ concentration during SAXS measurements. Hence, our experimental SAXS measurements strongly suggest that Codanin1$_{Cterm}$ occurs primarily as a monomer.

CDIN1 and Codanin1$_{Cterm}$ form a stable heterodimeric complex with nanomolar binding affinity

To reveal CDIN1-Codanin1$_{Cterm}$ interacting stoichiometry and complex composition, we used an electrophoretic mobility shift assay (EMSA) and size-exclusion chromatography (SEC) combined with SDS-PAGE (Fig. 2A, B). In the EMSA experiment, we allowed an increasing amount of CDIN1 to bind fluorescently labeled Codanin1$_{Cterm}$. We observed that the EMSA band corresponding to unbound Codanin1$_{Cterm}$ disappeared after reaching a 1:1 molar ratio CDIN1:Codanin1$_{Cterm}$, suggesting equimolar stoichiometry of complex formation. Additionally, we detected a small fraction of molecules arranged into a higher-order complex which could be a heterotetramer (Fig. 2A). Similarly, analytical SEC elution profiles revealed two clearly distinct peaks suggesting that CDIN1 and Codanin1$_{Cterm}$ form heterodimeric and heterotetrameric complexes together (Fig. 2B). Moreover, SDS-PAGE analysis of fractions corresponding to peak maxima revealed the comparable intensity of CDIN1 and Codanin1$_{Cterm}$ bands (see lanes 1 and 2 in the inset of Fig. 2B). The similar intensities of both proteins support the equimolar content of CDIN1 and Codanin1$_{Cterm}$ in the complex.

To further quantify how individual proteins arrange into a complex, we employed analytical ultracentrifugation (AUC). The sedimentation AUC profiles obtained (Fig. 2C) indicate that CDIN1 and Codanin1$_{Cterm}$ preferentially form heterodimeric and possibly tetrameric complexes. In the sedimentation profile of mixture CDIN1 and Codanin1$_{Cterm}$, we also detected a signal of free Codanin1$_{Cterm}$ due to its slight excess in the mixture. The sedimentation observations agree with previous SEC results, showing the preference for heterodimer formation and confirming a CDIN1-Codanin1$_{Cterm}$ complex ratio close to 1:1.

Next, we evaluated CDIN1-Codanin1$_{Cterm}$ mixture using SEC-MALS (Fig. 2D). The highest SEC signal corresponds to a molecular weight of 60 kDa, exactly matching the theoretical molecular weight of the CDIN1-Codanin1$_{Cterm}$ heterodimer (Fig. 2D). We also detected the broad SEC signal in the range of 112-130 kDa that most probably indicates a heterotetrameric complex of CDIN1 and Codanin1$_{Cterm}$.

Thus, combined separation approaches and scattering analyses revealed that CDIN1 and Codanin1$_{Cterm}$ assemble into equimolar heterodimers and heterotetramers.

To quantify the interaction between CDIN1 and Codanin1$_{Cterm}$, we carried out microscale thermophoresis (MST) and isothermal titration calorimetry (ITC) measurements. During MST analysis (Fig. 2E), we observed a change in fluorescence intensity of labeled Codanin1$_{Cterm}$ upon addition of CDIN1. The MST single-inflection sigmoidal curve fit revealed that CDIN1 binds Codanin1$_{Cterm}$ with a K_D value of 230 ± 50 nM. To validate the nanomolar binding affinity, we performed calorimetric titration with the complementary arrangement (Fig. 2F) - CDIN1 in the cell into which Codanin1$_{Cterm}$ was gradually injected. ITC thermograms provided a K_D value of 20 ± 4 nM for Codanin1$_{Cterm}$ binding to CDIN1. The partial oligomerization of individual proteins could explain why the stoichiometry obtained from ITC measurements is below the expected equimolar ratio. The difference in affinity values measured by MST and ITC might be caused by different temperatures, buffer conditions, and diverse method principles, fluorescence vs. heat exchange, used to measure binding affinity. Overall, we can confidently state that CDIN1 and Codanin1$_{Cterm}$ exhibit high affinity with K_D value in the 16-280 nanomolar range.

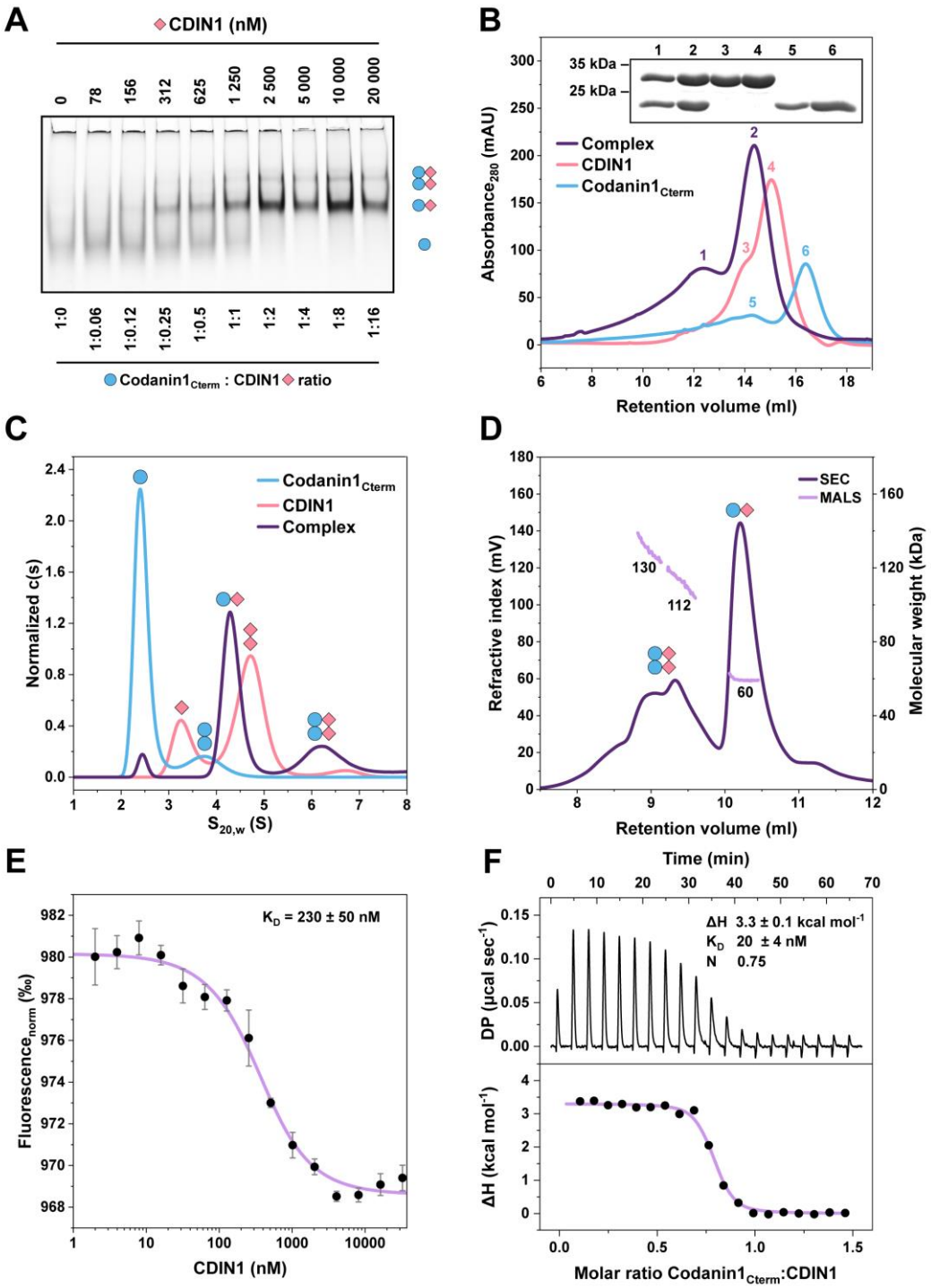


Figure 2. CDIN1 and Codanin1$_{Cterm}$ bind with equimolar stoichiometry and nanomolar affinity. **(A)** EMSA shows the binding of CDIN1 to fluorescently labeled Codanin1$_{Cterm}$ with highlighted molar ratios. Suggested protein arrangements are depicted as pictograms at band levels. **(B)** Size-exclusion chromatograms of individual and equimolarly incubated proteins demonstrates that CDIN1 and Codanin1$_{Cterm}$ form stable complexes. SDS-PAGE analysis of protein content in corresponding SEC fractions (inset). **(C)** Non-overlapping AUC profiles for individual proteins and complexes shows heterodimer formation. **(D)** SEC-MALS of CDIN1-Codanin1$_{Cterm}$ complex cross-validates heterodimer formation with corresponding molecular weight. **(E)** MST reveals the nanomolar affinity of CDIN1 binding to Codanin1$_{Cterm}$. **(F)** ITC titration of Codanin1$_{Cterm}$ to CDIN1 supports high mutual binding affinity.

CDA-I-related mutations residing in interacting regions disrupt CDIN1-Codanin1$_{Cterm}$ binding

To investigate the mutual interacting regions of CDIN1 and Codanin1$_{Cterm}$, we carried out hydrogen-deuterium exchange (HDX) coupled with mass spectrometry (MS). Based on tandem mass spectrometry detection connected with liquid chromatography (LC-MS/MS) bottom-up proteomics analysis of both proteins, the peptides covering whole protein sequences were identified. To evaluate HDX-MS data, we created the library from the peptides with high confidence, which covered 100% of Codanin1$_{Cterm}$, and 89% of CDIN1 (Fig. S6). We prepared samples of each protein alone or saturated with a two-fold excess of the interacting partner. Deuteration of individual ligand-free proteins or complexed proteins was quenched at 1 min, 10 min, 30 min, and 60 min as described previously[14,15]. To identify interacting regions, we quantified the relative deuterium uptake by comparing protein samples incubated for 10 min in H_2O and D_2O-based buffers. Deuterium uptake plots for other incubation times, which show no significant changes compared to the presented plot, can be found in the PRIDE repository[16]. The generated differential deuterium uptake plots displayed the difference in deuterium uptake between the bound and unbound states of the proteins, indicating regions that were less exposed to solvent upon binding (Fig. 3B, E). To interpret the data, we used modified Woods plots displaying the length and positions of residues of each analyzed peptide against their respective difference in relative deuterium uptake (ΔHDX)[17].

We identified two regions in CDIN1 that significantly prevented deuterium uptake upon binding (elevated ΔHDX): amino acids 63-98, and 143-172 (Fig. 3A, B). In Codanin1$_{Cterm}$, we identified three regions with suppressed deuteration (less solvent-exposed region), suggesting interaction sites of CDIN1 and Codanin1$_{Cterm}$ complex: amino acids 1028-1046, 1064-1087, and 1133-1170 (Fig. 3D, E).

Interestingly, we found that the interacting regions identified by HDX-MS contain mutations that have been described in CDA-I-suffering patients (Fig. 3A, D). Specifically, in interacting regions of CDIN1, we localized CDA-I-related mutations Y94C, Y94S, C156Y, and L178Q. In Codanin1$_{Cterm}$ interacting sequence, we found three disease-associated mutations – S1036F, R1042W, and D1043V.

To evaluate how introducing CDA-I-related mutations that reside in interacting regions affects the mutual binding of CDIN1 and full-length Codanin1, we employed site-directed mutagenesis followed by co-immunoprecipitation (Co-IP). We investigated the interaction of CDIN1 and Codanin1 using proteins expressed in HEK293T to ensure analysis of full-length post-translationally processed proteins in forms that occur in human cells. Importantly, our Co-IP results showed that both full-length Codanin1 and truncated Codanin1$_{Cterm}$ bind CDIN1, providing strong support for using truncated Codanin1$_{Cterm}$ in experiments describing interaction with CDIN1 (Fig. 3F, G). When we introduced a cluster of three closely located mutations S1036F-R1042W-D1043V (FWV) into Codanin1 or Codanin1$_{Cterm}$ (Fig. 3C), we observed no detectable interaction with CDIN1 (Fig. 3F), suggesting that the mutation cluster in Codanin1 prevents binding to CDIN1.

In the complementary experimental setup, we investigated how mutations in interacting regions of CDIN1 affect binding to Codanin1. Initially, we introduced two triplets of mutations Y94C-C156Y-L178Q (CYQ) and Y94S-C156Y-L178Q (SYQ) into CDIN1 (Fig. 3C). We found that introducing mutation triplets into CDIN1 aborted binding to Codanin1 (Fig. 3G).

To further identify which mutations are critical for CDIN1 binding to Codanin1, we prepared CDIN1 variants with single mutations. Our Co-IP experiments revealed that exchanging nonpolar tyrosine to polar cysteine or serine at CDIN1 position 94 in mutants Y94C and Y94S disrupted binding to Codanin1. Similarly, exchanging hydrophobic leucine for polar glutamine in the L178Q mutant of CDIN1 interrupted binding to Codanin1. On the contrary, introducing polar tyrosine instead of cysteine in C156Y mutant of CDIN1 did not affect Codanin1 binding (Fig. 3H, I). Thus, three of four mutations in the interacting region of CDIN1 are individually sufficient to abolish binding with Codanin1. Overall, we showed that CDA-I-related mutations in interacting regions CDIN1 and Codanin1 disturb the binding of the two proteins.

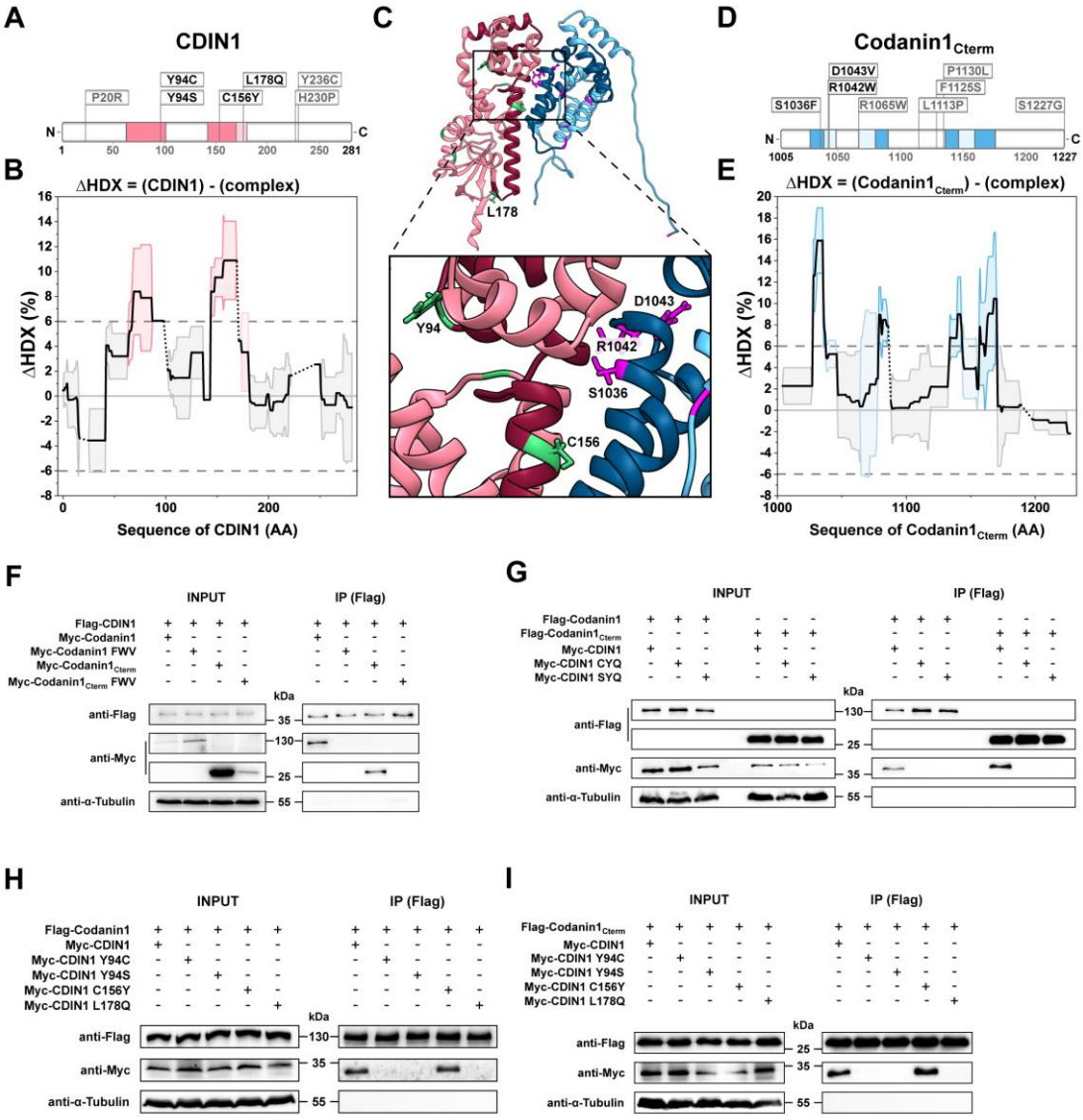


Figure 3. CDA-I-associated mutations in interacting regions disrupt CDIN1 and Codanin1 binding. **(A**, **D)** Scheme of CDIN1 and Codanin1$_{Cterm}$ highlighting CDA-I-associated mutations and interacting regions revealed by HDX-MS. **(B, E)** Significant sum differences in relative deuterium uptake (ΔHDX = CDIN1 - complex) for CDIN1 upon the addition of two-fold molar excess of Codanin1$_{Cterm}$ and differences in relative deuterium uptake (ΔHDX = Codanin1$_{Cterm}$ - complex) for Codanin1$_{Cterm}$ upon the addition of two-fold molar excess of CDIN1 after 10 minutes of deuteration, respectively. Highlighted colored regions represent interacting sites with mean value (dark colored) or standard deviation (light colored) exceeding the 6% deuteration threshold. **(C)** Heterodimer CDIN1-Codanin1$_{Cterm}$ structure predicted by AlphaFold-Multimer. Interacting regions determined by HDX-MS are highlighted by darker colors – red for CDIN1 and blue for Codanin1$_{Cterm}$. CDA-I-associated mutations in CDIN1 are shown in green; in Codanin1$_{Cterm}$ in magenta. The mutated amino acids in interacting regions are displayed as sticks and numbered. **(F, G)** Complementary Co-IP analyses using wild type and triple-mutated Codanin1 and CDIN1 proteins expressed in HEK293T cells show that simultaneous introduction of CDA-I-associated mutations in either Codanin1 or CDIN1 interacting regions disrupts CDIN1-Codanin1 binding. **(H, I)** Co-IP analyses using proteins expressed in HEK293T cells identify individual CDIN1 mutations associated with CDA-I that are critical for disrupting the binding to full-length Codanin1 or Codanin1$_{Cterm}$.

Discussion

In this work, we report the first low-resolution structures of CDIN1 and Codanin1, two proteins associated with CDA-I (Fig. 1). We found that CDIN1 and Codanin1$_{Cterm}$ bind with nanomolar binding affinity and form stable complexes (Fig. 2). We defined interacting regions of CDIN1 and Codanin1$_{Cterm}$ and identified CDA-I-related mutations disrupting CDIN1-Codanin1 binding (Fig. 3).

Our structural studies of individual proteins have indicated that CDIN1 preferentially forms dimers. The prevalent homodimeric state of CDIN1 has been detected by three different methods in a wide range of concentrations – AUC (0.6 mg ml^{-1}), SEC-MALS (2 mg ml^{-1}), and SEC-SAXS (10 mg ml^{-1}), suggesting that CDIN1 occurs as stable concentration-independent dimers. We used the same methods to reveal that Codanin1$_{Cterm}$ occurs preferentially as a monomer. However, with increasing concentration of Codanin1$_{Cterm}$, we observed a higher amount of non-monomeric fraction suggesting concentration dependence of monomeric-multimeric equilibrium of Codanin1$_{Cterm}$. Furthermore, we reconstructed SAXS structural envelopes accommodating dimeric CDIN1 and monomeric Codanin1$_{Cterm}$ projections. We should nevertheless consider the limitations of our findings for Codanin1$_{Cterm}$, as we analyzed the isolated C-terminal region of Codanin1 that could be affected by the other regions of full-length Codanin1.

We have demonstrated that CDIN1 and Codanin1$_{Cterm}$ bind in vitro. The direct binding of CDIN1 and Codanin1 is consistent with previous immunoprecipitation studies[2,10]. The nanomolar affinity of CDIN1 and Codanin1 determined is comparable with the affinity of the RSC (remodeling the structure of chromatin) complex[18,19]. We propose that CDIN1 and Codanin1$_{Cterm}$ form complexes in near equimolar ratio based on stoichiometry estimated from EMSA (Fig. 2A) and the densitometry analysis of PAGE profiles corresponding to SEC peaks (Fig. 2A, B). Recorded SEC profiles contained signals at a high-molecular weight range, suggesting that both proteins form higher-order oligomers (Fig. 1).

After successfully defining interacting regions by HDX-MS, we confirmed that CDA-I-associated mutations in CDIN1 and Codanin1 interacting regions disrupt their mutual binding. We identified three mutations in CDIN1 that are related to disease and individually diminish CDIN1's ability to bind to Codanin1. Furthermore, we identified a mutation triplet in Codanin1 that prevents CDIN1 binding.

Codanin1 uses its N-terminal B-domain to bind ASF1, sequestering ASF1 in the cytoplasm. In addition, Codanin1 interacts with ASF1 using the same motif on ASF1 that is responsible for binding with two other downstream histone chaperones, CAF-1 and HIRA. Codanin1 thus suppresses ASF1-mediated histone deposition[7]. Interestingly, a few CDA-I-associated mutations, such as R714W and R1042W, might compromise the interaction between Codanin1 and ASF1[7]. Sharing the C-terminal R1042 site, crucial for CDIN1 binding, suggests that CDIN1 and ASF1 may partially compete for the same binding region on Codanin1. Alternatively, a trimeric complex might be formed if we consider multiple binding sites on Codanin1.

Our results raise the fundamental question: what is the molecular function of CDIN1? We speculate that CDIN1, via its binding to the C-terminal part of Codanin1, may induce allosteric changes, possibly by arranging adjacent intrinsically disordered regions of Codanin1 into functional structures. The proper CDIN1-mediated arrangement of Codanin1 might be critical for the effective regulation of histone trafficking essential for proper chromatin formation. The disturbance of tightly controlled histone depositions causes spongy chromatin and internuclear bridges in erythroblasts – the typical hallmarks of CDA-I disease (Fig. 4). Further research using erythroblast model cell lines should be carried out to determine if mutations that disrupt CDIN1-Codanin1 binding also result in a phenotype characteristic of CDA-I.

In summary, our research offers a framework for the structural understanding of how CDIN1 and Codanin1 interact and thus contribute to chromatin remodeling that must be tightly handled during DNA replication and gene transcription. Here we quantified the interaction of CDIN1 and Codanin1 and defined interacting regions. We identified mutations in CDIN1 and Codanin1 disturbing the mutual interaction of proteins tightly associated with CDA-I disease progression. Further, our finding of structural arrangements of CDIN1 and Codanin1 is an important step towards elucidating the molecular origin of the disease. The results presented here improve our understanding of the molecular mechanisms of erythropoiesis, whose disturbance leads to CDA-I.

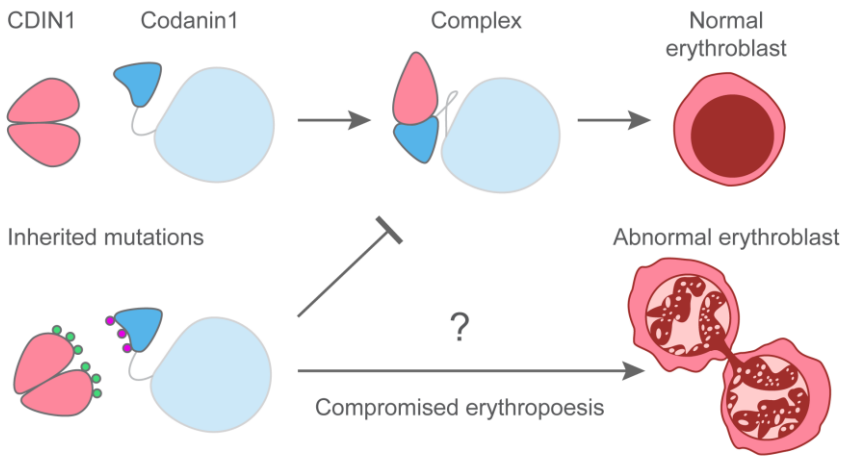


Figure 4: A potential role of CDIN1 and Codanin1 interaction in CDA-I progression.
We propose a hypothetical model in which CDIN1-Codanin1 interaction is critical for a healthy development of erythroblasts. When wild-type CDIN1 binds Codanin1, common erythropoiesis and the development of normal erythroblasts occur. CDA-I-associated mutations disrupt the mutual binding of CDIN1 and Codanin1. We hypothesize that preventing CDIN1 and Codanin1 interaction affects regulatory functions of Codanin1 in histone deposition, compromising erythropoiesis and manifesting by abnormal erythroblasts in the process.

Materials and methods

Origin of CDIN1 and Codanin1 coding sequences

The total RNA was extracted from 3 x 10^6 HEK293T cells using the RNeasy Mini Kit (QIAGEN). RNA served as a template for subsequent reverse transcription into cDNA by iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad). Specific primers (cDNA_CDIN1_Fw ATGATACTGACCAAAGCTCAGTAC and cDNA_CDIN1_Rv TCAAGCTATGCTGTGGCATAA) were designed to amplify CDIN1 coding sequence using proofreading Q5 High-Fidelity DNA Polymerase (New England Biolabs). The PCR product was purified by GenElute PCR Clean-Up Kit (Sigma-Aldrich) and cloned into pCR-Blunt II-TOPO vector by Zero Blunt TOPO PCR Cloning Kit (Invitrogen). All procedures were done according to protocols from the manufacturers. Plasmids containing the whole human Codanin1 sequence Topo_pCR2_CDAN1 and Codanin1 sequence optimized for bacterial expression pHAT4_CDAN1 were kindly provided by Prof. Anja Groth (University of Copenhagen).

Cloning, expression, and purification of CDIN1 and Codanin1 variants

NEBuilder® HiFi DNA Assembly kit (New England Biolabs) was used to clone CDIN1 and Codanin1$_{Cterm}$ coding sequences into HindIII, XhoI restriction site of plasmid pl21 (TriAltus) that is suitable for Im7/CL7 ultra-high affinity chromatography purification from bacteria. Similarly, Flag- and Myc-tagged variants of CDIN1, Codanin1, and Codanin1$_{Cterm}$ were cloned into NotI, SalI restriction site of plasmid pHAGE_puro (Addgene #118692) that allows expression in mammalian systems. Cloning was performed according to the recommendation from the manufacturer. All constructs were transformed into *E. coli* strain One Shot™ TOP10 (Invitrogen) for plasmid amplification or BL21-CodonPlus (DE3)-RIPL Competent Cells (Agilent) for protein expression in bacteria. Successful cloning was confirmed by Sanger sequencing.

To express proteins, BL21(DE3) RIPL cell cultures harboring pl21_CDIN1 and pl21_Codanin1$_{Cterm}$ were grown in Terrific Broth (TB) medium containing 50 µg ml^{-1} kanamycin and 34 µg ml^{-1} chloramphenicol in orbital shaker New Brunswick™ Innova® 43/43R (New Brunswick) at 160 rpm and 37 °C. When OD_{600} reached a value between 0.6-0.8, the cell cultures were cooled down to 20 °C, and protein expression was induced by IPTG addition to the final concentration of 0.1 mM. After induction, the cell cultures were cultured O/N and collected the next day by centrifugation (8 000 g, 8 min, 4 °C) and stored at -80 °C.

The pellet was resuspended in lysis buffer (20 mM Tris-HCl, 1000 mM NaCl, 5% glycerol, 0.05% Triton X-100, pH 8.0 with the addition of 1 mM DTT, protease inhibitor cocktail cOmplete tablets EDTA-free (Roche), and 0.4 mg ml^{-1} of lysozyme before use. The cell suspension was sonicated on ice with amplitude settings 30 for 4 min of process time, where the pulse-on and pulse-off intervals were 1 and 3 seconds, respectively (Misonix S-4000). The cell lysate was centrifugated at 60 000 g, 4 °C for 30 minutes and filtered through 0.45 µm Sterivex™ (Merck Millipore). Supernatants of CDIN1 and Codanin1$_{Cterm}$ were purified by a three-step protocol utilizing CL7/Im7 Affinity Tag System (TriAltus), His-Trap HP (Cytiva), and size-exclusion chromatography (Cytiva). Firstly, filtered supernatant was loaded onto the Im7 FPLC column (5 ml), and then the column was washed by five column volumes of Im7 buffer (20 mM Tris-HCl, 500 mM NaCl, 5% glycerol, pH 8.0). For protein elution, the proteolytic cleavage by home-made His-tagged Ulp1 protease diluted in Im7 buffer (final concentration 0.1 mg ml^{-1}) with the addition of 1 mM DTT was used. An elution sample from Im7 FPLC column was loaded onto the His-Trap HP column to clear the sample from residual His-tagged Ulp1. In the case of CDIN1 and Codanin1$_{Cterm}$, non-specific binding to His-Trap HP column was observed. To reduce nonspecific binding 5-10% gradient of His-Trap buffer (20 mM Tris-HCl, 500 mM NaCl, 500 mM imidazole, 5% glycerol, pH 8.0) in Im7 buffer was used to elute proteins. For the buffer exchange and final exclusion of possible impurities, CDIN1 and Codanin1$_{Cterm}$ were loaded onto the HiLoad 16/600

column containing Superdex 200 pg (Cytiva) and resolved using final buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0) that was used in most experiments afterward. In the last step, proteins were concentrated by ultrafiltration on 10K Amicon® Ultra Centrifugal Filters (Merck Millipore), aliquoted, snap-freezed in liquid nitrogen, and stored at -80 °C. Protein identity and purity were analyzed by MALDI-TOF mass spectrometry (Proteomics Core Facility, CEITEC MU, Brno) and 12% SDS-PAGE with consequent Coomassie Brilliant Blue G250 staining for the following densitometry analysis[6].

Electrophoretic mobility shift assay (EMSA)

In EMSA experiment, Codanin1$_{Cterm}$ (1.25 µM) fluorescently labeled by Alexa Fluor 488 was incubated with increasing amounts of CDIN1 (0-20 µM, 2-fold dilution line) in a final buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0) for 15 min at room temperature. The binding reactions were resolved on native polyacrylamide gels (10% PAGE, 0.5×TBE). The runs were performed at 10 V cm^{-1} (Mini-PROTEAN spacer plate, Bio-Rad) for 45 minutes (or until loading dye reached the bottom of the gel) at 4 °C, and gels were visualized using a 473 nm laser and bandpass filter (510-530 nm) on a Typhoon™ FLA 9500 biomolecular imager (GE Healthcare Life Sciences).

Size-exclusion chromatography (SEC)

Protein samples for size-exclusion chromatography were filtered through Nanosep filters with 0.45 µm wwPTFE membrane (Pall) to avoid the presence of aggregates after the freeze and thaw cycle. CDIN1 (1 mg) and Codanin1$_{Cterm}$ (0.75 mg) proteins were analyzed separately or as a complex preincubated for 15 minutes in a molar ratio 1:1, where the amount of the proteins was halved (0.5 mg and 0.375 mg) to achieve comparable signal levels.

Chromatographic separation was performed on 10/300 GL Superdex 200 column (Cytiva) using a final buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0) as a mobile phase and flow rate 0.3 ml min^{-1}. Fractions corresponding to the peaks from the chromatogram were collected and analyzed on 12% SDS-PAGE.

Microscale thermophoresis (MST)

MST measurements were performed using a NanoTemper Monolith NT.115 instrument (NanoTemper Technologies). Fluorescently labeled Codanin1$_{Cterm}$ Alexa Fluor 488 (250 nM) was incubated for 15 min on ice with different concentrations of CDIN1 (0-50 µM, 2-fold dilution line), in a buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0). The samples were loaded into premium capillaries. MST measurements were performed at 20 °C, 40% blue-laser power, and 45% LED power (MST power). The laser-on and laser-off intervals were 20 and 5 seconds, respectively. MO.Affinity Analysis v2.2.4 software was used to fit the data by Hill equation and to determine apparent K_D values from MST traces at 2.5 s time point after the infra-red laser irradiation. All measurements were performed at least three times for two independently prepared sample sets.

Isothermal titration calorimetry (ITC)

ITC experiments were carried out on a VP-ITC instrument (Microcal, GE Healthcare) at 15 °C. Proteins were dialyzed into the final buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0) and degassed before the experiment. The ITC cell (1423 µl) was filled with CDIN1 (11 µM). Codanin1$_{Cterm}$ (110 µM) was added in 20 consecutive titrations of 10 µl in 5-minute intervals, with a stirring rate of 240 rpm. Experimental data were analyzed in OriginPro 2022 version 9.9.0.225 software (OriginLab

Corporation) using a one-site binding model to fit a theoretical titration curve. Binding enthalpy (ΔH), binding constant (K_D), and reaction stoichiometry (n) were obtained from the fit.

SEC Multiangle Light Scattering (MALS)

Proteins CDIN1 (2 mg ml^{-1}) and Codanin1$_{Cterm}$ (2 mg ml^{-1}) or complex CDIN1-Codanin1$_{Cterm}$ were preincubated and dialyzed before the measurement in a buffer 20 mM Tris-HCl, 150 mM NaCl, pH 8.0 and then loaded onto an OmniSEC instrument (Malvern Panalytical) equipped with a 13-ml Zenix SEC 300 column (Sepax Technologies) previously equilibrated in the same buffer coupled to a Wyatt miniDAWN TREOS and Optilab T-rEX differential refractive index detector. Molar mass was calculated from the Raleigh ratio based on the multiangle (static) light scattering and protein concentration from the change in refractive index (dn/dc = 0.185). Analysis was performed using OMNISEC software 11.21 (Malvern Panalytical).

Analytical ultracentrifugation (AUC)

Analytical ultracentrifugation experiments were performed using ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter) equipped with an An-60 Ti rotor at 20 °C and 50 000 rpm. Sedimentation velocity (SV) experiment was performed in 12 mm titanium double-sector centerpiece cells (Nanolytics Instruments) loaded with 380 µl of both protein sample and buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8.0). Scans were collected at 280 nm in 5 minutes intervals and 0.003 cm spatial resolution in continuous scan mode. Protein samples were diluted in reference buffer to the final concentration of 24 µM CDIN1, 24 µM Codanin1$_{Cterm}$ and complex CDIN1-Codanin1$_{Cterm}$ in a ratio 1:1, where the concentration of the proteins was lowered to half (12 µM for both) to achieve similar signal levels.

The partial specific volume of the protein and the solvent density and viscosity were calculated from the amino acid sequence and buffer composition, respectively, by the Sednterp software (http://bitcwiki.sr.unh.edu). The SV data corresponding to the complete sample sedimentation were fitted using the continuous c(s) distribution model in Sedfit v15.01c[20]. The M_W of particles was estimated based on the Svedberg equation, and c(s) distributions were normalized and plotted using the GUSSI software version 1[20].

SEC Small-Angle X-ray Scattering (SAXS)

The SAXS experiments were performed using a SAXSpoint 2.0 (Anton Paar, Graz, Austria) equipped with a MetalJet C2+ X-ray source (Excillum, Stockholm, Sweden) and an Eiger R 1M detector (Dectris, Baden, Switzerland). CDIN1 and Codanin1$_{Cterm}$ were diluted in 20 mM Tris HCl, 150 mM NaCl, pH 8.0 buffer. The data were collected at 20 °C for the buffer and protein samples at a 10 mg ml^{-1} of loading concentration. CDIN1 data in selected intervals (frames 504-615) were solvent-subtracted in ATSAS Chromixs and then characterized using the ATSAS suite[13]. The Codanin1$_{Cterm}$ data in selected intervals (frames for first peak 605-643 and for second peak 762-817) were processed equally. The R_g values and P(r) distribution function calculations were performed using GNOM and PRIMUS[21]. The ab initio reconstruction was performed with DAMMIF[22], aligned and averaged with DAMAVER[23] to generate the most probable model, and refined with DAMMIN[24] to generate the final SAXS envelope. The agreement between the SAXS data and AlphaFold predicted structures of CDIN1 and Codanin1$_{Cterm}$ was evaluated using CRYSOL and DAMMIF generated SAXS envelope. Experimental SAXS data and derived models of both CDIN1 and Codanin1$_{Cterm}$ have been deposited in the Small Angle Scattering Biological DataBank (SASBDB)[25].

Hydrogen-deuterium exchange mass spectrometry (HDX-MS)

CDIN1 and Codanin1$_{Cterm}$ samples were diluted to the final concentration of 2 µM in H_2O-based buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8.0) to prepare undeuterated controls for peptide mapping. Free deuterated samples CDIN1 and Codanin1$_{Cterm}$ or their complex in the molar ratios 1:2 and 2:1 (CDIN1-Codanin1$_{Cterm}$) were preincubated for 30 minutes at room temperature and then diluted in D_2O-based buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8.0). To monitor hydrogen-deuterium exchange (HDX), samples were incubated at room temperature and quenched after 1 minute, 10 minutes, 30 minutes, or 1 hour by the addition of 0.5 M TCEP-HCl, 4 M urea, 1 M glycine (pH 2.3) containing pepsin in final concentration 0.04 mg ml^{1} followed by 3 min incubation and rapid freezing in liquid nitrogen. HDX data collected at the 10 min interval are presented because HDX results obtained after all measured time intervals were similar.

Each frozen sample was thawed and injected into the LC-system (UltiMate 3000 RSLCnano, Thermo Scientific), where the protein was digested within a dual-protease enzymatic column (Nepenthesin-1 and pepsin, 15 µl bed volume, Affipro s.r.o., CZ). Peptides were trapped and desalted on-line on a peptide microtrap column (Michrom Bioresources, Auburn, CA) for 3 min. Both digestion and desalting were performed in loading buffer (2% acetonitrile, 0.05% trifluoroacetic acid) at a flow rate of 100 µl min^{-1}. Next, the peptides were eluted onto an analytical column (Jupiter C18, 0.5 x 50 mm, 5 µm, 300 Å, Phenomenex, CA) and separated by a 29 min linear gradient elution starting with 10% buffer B (80% acetonitrile in 0.08% formic acid) in buffer A (0.1% formic acid) and rising to 40% buffer B at a flow rate of 50 µl min^{-1}. The dual-protease, trap, and analytical columns were kept at 1.5 °C during the whole procedure.

Mass spectrometric analysis was carried out using an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) with ESI ionization connected online to a robotic system based on the HTS-XT platform (CTC Analytics, Zwingen, Switzerland). The instrument was operated in a data-dependent mode for peptide mapping (LC-MS/MS). Each MS scan was followed by MS/MS scans of the three most intensive ions from both collision-induced dissociation (CID) and higher-energy C-trap dissociation (HCD) fragmentation spectra. Tandem mass spectra were searched using SequestHT against the cRap protein database (ftp://ftp.thegpm.org/fasta/cRAP) containing the sequences of CDIN1 and Codanin1$_{Cterm}$ recombinant proteins with the following search settings: mass tolerance for precursor ions of 10 ppm, mass tolerance for fragment ions of 0.6 Da, no enzyme specificity, two maximum missed cleavage sites and no-fixed or variable modifications. The false discovery rate at the peptide identification level was 1%. Sequence coverage was analyzed with Proteome Discoverer version 1.4 (Thermo Fisher Scientific) and graphically visualized with the MS Tools application[26].

The analysis of deuterated samples was done in LC-MS mode with ion detection in the orbital ion trap. The MS raw files, together with the list of peptides (peptide pool) identified with high confidence characterized by requested parameters (amino acid sequence of each peptide, its retention time, XCorr, and ion charge), were processed using HDExaminer version 2.5 (Sierra Analytics, Modesto, CA). The software analyzed protein and peptides behavior, created the uptake plots showing sum of differences in relative deuterium uptake (ΔHDX = free protein - complex) describing peptide deuteration over time with a calculated confidence level (high and medium confidence are accepted, low confidence is rejected). Each of the accepted peptides with high confidence level was mapped to the amino acid sequences of the analysed proteins via the following procedure. Each residue was assigned the uptake data from any peptide solved with high confidence[14]. Low and medium confidence peptides were rejected. The final uptake value (expressed as % of deuteration) assigned to each amino acid corresponded to the average of all assigned values for its position.

To check the suitability of our LC-MS system, HDX-MS analysis of a standard protein, myoglobin, was performed[15]. Reproducible deuterium recovery, calculation of the back-exchange level, and other recommendations published in Nature Methods[27] were conducted. The mass spectrometry

proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE repository[16] with the dataset identifier PXD037661.

Co-immunoprecipitation (Co-IP)

For Co-IP, HEK293T cells were cultivated in 37 °C and 5% CO_2 using 12 ml of 1x Dulbecco's Modified Eagle Medium per each 10-cm Petri dish – DMEM, high glucose, pyruvate supplemented, 4 mM L-glutamine, 1x MEM non-essential amino acids solution, 1x Penicillin-Streptomycin (all Gibco) and 10% FBS (Capricorn). When HEK293T cells reached almost full confluency, cells were washed by 5 ml of 1x PBS and detached by addition of 1 ml Trypsin-EDTA (0.25%), phenol red (both Gibco) for 1 minute, diluted in DMEM media, and split using one third of the cells per new 10-cm Petri dish. Next day, HEK293T were co-transfected with 10 µg of the plasmids containing Flag- and Myc-tagged protein constructs using polyethylenimine (PEI) on 10-cm Petri dishes.

The cells were harvested 48 hours after transfection, washed once with PBS, and lysed for 1.5 hours rotating at 4 °C in 500 µl of RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, pH 7.4, with the addition of 1 mM DTT, 10 mM NaF, and protease inhibitor cocktail cOmplete tablets EDTA-free (Roche) right before use). After incubation, lysates were spun at 20 000 g for 15 min to pellet debris. Supernatants were mixed with 40 µl of Anti-FLAG® M2 Magnetic Beads (Millipore) and incubated for 3 hours while rotating at 4 °C. The beads were washed three times in RIPA buffer without additives (50 mM Tris-HCl, 150 mM NaCl, 0.5% Triton X-100, 1 mM EDTA, pH 7.4) and three times with wash buffer (50 mM Tris, 150 mM NaCl, pH 7.4). For the following elution, the beads were resuspended in 60 µl of elution buffer (50 mM Tris, 150 mM NaCl, pH 7.4 containing 3x Flag peptide in final concentration 150 µg ml^{-1} (Sigma-Aldrich) for 1.5 hours at 4 °C. Input supernatant and elution samples were separated on 12% SDS-PAGE and transferred onto a membrane (Cytiva, Amersham™ Protran®).

The following analysis was done by western blot using Monoclonal ANTI-FLAG® M2-Peroxidase (HRP) antibody (Sigma-Aldrich) or Monoclonal Anti-Myc tag antibody [9E10] (Abcam) produced in mice. Flag antibody was diluted in a ratio of 1:3 000, Myc antibody was diluted in a ratio of 1:1 000 in blocking buffer (3.5% milk in TBS-T), and incubated O/N at 4 °C. Anti-Flag antibody was detected immediately after TBS-T wash (1x TBS, 0.1% Tween20). To detect anti-Myc antibody, the membrane was washed three times by TBS-T, incubated with secondary Anti-Mouse IgG (Fc specific) HRP (SigmaAldrich) diluted in a ratio of 1:5 000 for 1 hour at room temperature, and then washed three times again before detection by SuperSignal™ West Femto Maximum Sensitivity Substrate kit (Thermo Scientific) on Fusion FX instrument (Vilber).

Circular dichroism (CD)

Far-UV CD measurements were performed on a J-815 spectrometer (Jasco) at 20 °C in 1-mm Quartz cuvette (Hellma Analytics). CD spectra of 0.2 mg ml^{-1} CDIN1 and 0.2 mg ml^{-1} $Codanin1_{Cterm}$ proteins in a buffer (20 mM NaH_2PO_4, 150 mM NaF, pH 8.0) were acquired in the wavelength range of 190-250 nm with a 1 nm step at a scanning speed of 100 nm min^{-1}. Each spectrum represents an average of ten accumulations. Subsequently, the buffer signal was subtracted, and data were converted from circular dichroism units to mean residue molar ellipticity (MRE) to account for precise protein concentration. The presence of secondary structural elements was evaluated using BeStSel software [28].

Differential scanning fluorimetry (nanoDSF)

Protein thermal stability was determined by differential scanning fluorimetry (nanoDSF) on Prometheus NT.48 instrument (Nanotemper Technologies) in nanoDSF grade standard capillaries

filled with the samples. The measurements were performed in triplicates using 2 mg ml^{-1} CDIN1 and 1.5 mg ml^{-1} $Codanin1_{Cterm}$ in a buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0), the temperature in the range of 20-80 °C at a heating rate of 1 °C min^{-1} and at excitation power 10%. Protein unfolding was monitored by fluorescence intensity measured at 330 and 350 nm. Subsequently, the melting temperature (T_m) was determined from the first derivative of fluorescence ratio (330/350)[20].

Dynamic light scattering (DLS)

The tendency of the protein to form aggregates was determined by dynamic light scattering (DLS) in the Delsa Max Core (Beckman Coulter). Prior to measurement, 4 mg ml^{-1} CDIN1 and 4 mg ml^{-1} $Codanin1_{Cterm}$ in a buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0) were centrifuged at 12 000 g, RT for 8 min. Twenty scans of 10 s data acquisition were averaged to examine the sample homogeneity. Data were collected and processed by software provided by Beckman Coulter. The presence of aggregates was determined based on a regularization fit of obtained autocorrelation functions of scattered light. The data were evaluated as the qualitative analysis of the aggregate's content (the presence of peaks of *R*h > 100 nm) for intensity-based data.

Native PAGE

CDIN1 and $Codanin1_{Cterm}$ (4 µg) in a final buffer (20 mM Tris-HCl, 150 mM NaCl, 5% glycerol, pH 8.0) were mixed with reducing and non-reducing native loading dyes. Afterwards, 12% native PAGE was performed in 0.5x TBE buffer. Results were compared with the samples prepared in denaturing conditions with reducing and non-reducing SDS containing loading dyes that were run on standard 12% SDS-PAGE in TANK buffer.

Competing interests

None declared.

Acknowledgments and Funding

The Czech Science Foundation (GA23-05241S to C.H.) has primarily supported this research. The research has been carried out with institutional support of the Institute of Biophysics of the Czech Academy of Sciences (68081707 to M.S., P.V, and C.H.).

CIISB, Instruct-CZ Centre of Instruct-ERIC EU consortium, funded by MEYS CR infrastructure project LM2023042 and European Regional Development Fund-Project "UP CIISB" No. CZ.02.1.01/0.0/0.0/18_046/0015974, is gratefully acknowledged for the financial support of the measurements at the following core facilities: CEITEC – Proteomics, Biomolecular Interactions and Crystallography, CMS-Biocev – Biophysical techniques, Crystallization, Diffraction, and Structural mass spectrometry.

The research has been supported by project National Institute for Cancer Research (Programme EXCELES, ID Project No. LX22NPO5102) – Funded by the European Union – Next Generation EU and by the Ministry of Health, Czech Republic – Conceptual Development of Research Organization, MH CZ - DRO (MMCI, 00209805). T.B. is supported by Brno Ph.D. Talent Scholarship – funded by Brno City Municipality.

We recognize Agnel Sfeir for her brilliant igniting ideas and wholehearted support on both scientific and personal levels. We thank Samuel Tremblay-Belzile for his critical reading and editing of the manuscript.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE repository with the dataset identifier PXD037661 with access available upon request.

The small-angle X-ray scattering data have been deposited to Small-angle scattering biological data bank (SASBDB)[25]: CDIN1, Codanin1$_{Cterm}$ second peak, and Codanin1$_{Cterm}$ first peak.

References

1. Iolascon, A., Andolfo, I. & Russo, R. Congenital dyserythropoietic anemias. *Blood* **136**, 1274-1283 (2020).
2. Olijnik, A.A. et al. Genetic and functional insights into CDA-I prevalence and pathogenesis. *Journal of Medical Genetics* **58**, 185-195 (2021).
3. Wickramasinghe, S.N. & Wood, W.G. Advances in the understanding of the congenital dyserythropoietic anaemias. *British Journal of Haematology* **131**, 431-446 (2005).
4. Swickley, G. et al. Characterization of the interactions between Codanin-1 and C15Orf41, two proteins implicated in congenital dyserythropoietic anemia type I disease. *Molecular and Cell Biology* **21**(2020).
5. Dgany, O. et al. Congenital Dyserythropoietic Anemia Type I Is Caused by Mutations in Codanin-1. *The American Journal of Human Genetics* **71**, 1467-1474 (2002).
6. Renella, R. et al. Codanin-1 mutations in congenital dyserythropoietic anemia type 1 affect HP1α localization in erythroblasts. *Blood* **117**, 6928-6938 (2011).
7. Ask, K. et al. Codanin-1, mutated in the anaemic disease CDAI, regulates Asf1 function in S-phase histone supply. *The EMBO Journal* **31**, 2013-2023 (2012).
8. Babbs, C. et al. Homozygous mutations in a predicted endonuclease are a novel cause of congenital dyserythropoietic anemia type I. *Haematologica* **98**, 1383-1387 (2013).
9. Su, A.I. et al. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 6062-6067 (2004).
10. Shroff, M., Knebel, A., Toth, R. & Rouse, J. A complex comprising C15ORF41 and Codanin-1: the products of two genes mutated in congenital dyserythropoietic anaemia type I (CDA-I). *Biochemical Journal* **477**, 1893-1905 (2020).
11. Roy, N.B.A. & Babbs, C. The pathogenesis, diagnosis and management of congenital dyserythropoietic anaemia type I. *British Journal of Haematology* **185**, 436-449 (2019).
12. Manalastas-Cantos, K. et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *Journal of Applied Crystallography* **54**, 343-355 (2021).
13. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *Journal of Applied Crystallography* **50**, 1212-1225 (2017).
14. Schenkmayerova, A. et al. Engineering the protein dynamics of an ancestral luciferase. *Nature Communications* **12**, 3616 (2021).
15. Uhrik, L. et al. Hydrogen deuterium exchange mass spectrometry identifies the dominant paratope in CD20 antigen binding to the NCD1.2 monoclonal antibody. *Biochemical Journal* **478**, 99-120 (2021).
16. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Research* **47**, D442-d450 (2019).
17. Ahdash, Z. et al. HDX-MS reveals nucleotide-dependent, anti-correlated opening and closure of SecA and SecY channels of the bacterial translocon. *eLife* **8**, e47402 (2019).
18. Jain, N. et al. Interaction of RSC Chromatin Remodeling Complex with Nucleosomes Is Modulated by H3 K14 Acetylation and H2B SUMOylation In Vivo. *iScience* **23**, 101292 (2020).
19. Chatterjee, N. et al. Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms. *Nucleic Acids Research* **39**, 8378-8391 (2011).
20. Kozeleková, A. et al. Phosphorylated and Phosphomimicking Variants May Differ—A Case Study of 14-3-3 Protein. *Frontiers in Chemistry* **10**(2022).
21. Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography* **25**, 495-503 (1992).
22. Franke, D. & Svergun, D.I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **42**, 342-346 (2009).
23. Volkov, V.V. & Svergun, D.I. Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **36**, 860-864 (2003).

24. Svergun, D.I. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical Journal* **76**, 2879-2886 (1999).
25. Kikhney, A.G., Borges, C.R., Molodenskiy, D.S., Jeffries, C.M. & Svergun, D.I. SASBDB: Towards an automatically curated and validated repository for biological scattering data. *Protein Science* **29**, 66-75 (2020).
26. Kavan, D. & Man, P. MSTools—Web based application for visualization and presentation of HXMS data. *International Journal of Mass Spectrometry* **302**, 53-58 (2011).
27. Masson, G.R. et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods* **16**, 595-602 (2019).
28. Micsonai, A. et al. BeStSel: webserver for secondary structure and fold prediction for protein CD spectroscopy. *Nucleic Acids Research* **50**, W90-W98 (2022).
29. Drew, E.D. & Janes, R.W. PDBMD2CD: providing predicted protein circular dichroism spectra from multiple molecular dynamics-generated protein structures. *Nucleic Acids Research* **48**, W17-W24 (2020).