The dynamic properties of a nuclear coactivator binding domain are evolutionarily conserved

Elin Karlsson[1], Frieda A. Sorgenfrei[1,†], Eva Andersson[1], Jakob Dogan[1], Per Jemth[1,*], and Celestine N. Chi[1,2,*]

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden.

[2]Department of Pharmaceutical Biosciences, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden

[†]Present address: acib GmbH, Krenngasse 37, 8010 Graz c/o University of Graz, Institute of Chemistry, NAWI Graz, BioTechMed Graz, Heinrichstrasse 28, 8010 Graz, Austria

*Correspondence to

Celestine Chi, Celestine.Chi@imbim.uu.se

Per Jemth, Per.Jemth@imbim.uu.se

ABSTRACT

Evolution of proteins is constrained by their structure and function. While there is a consensus that the plasticity of intrinsically disordered proteins relaxes the structural constraints on evolution there is a paucity of data on the molecular details of these processes. The Nuclear co-activator binding domain (NCBD) from CREB-binding protein is a protein-protein interaction domain, which contains a hydrophobic core but is verging on being intrinsically disordered. These highly dynamic 'borderline' properties of NCBD makes it an interesting model system for evolutionary structure-function investigation. We have here compared the structure and biophysical properties of an ancient version of NCBD present in a bilaterian animal ancestor living around 600 million years ago with extant human NCBD. Using a combination of NMR spectroscopy, circular dichroism and kinetic methods we show that NCBD has retained its structure and dynamic biophysical properties in the ligand-free state in the evolutionary lineage leading from the bilaterian ancestor to humans. Our findings suggest that the dynamic properties of NCBD are subject to positive selection and thus important for its function, which includes mediating several distinct protein-protein interactions.

Introduction

The non-covalent chemistry of polypeptide chains gives rise to a wide spectrum of structure, stability and dynamics among proteins, with fully disordered proteins at one end of the scale and well-folded globular proteins at the other one [1]. It is becoming clear that dynamic properties *per se* are important for protein function but it is not trivial to prove that the dynamics are essential for function and not merely a general property of proteins. The Nuclear Coactivator Binding Domain (NCBD) is a small (approximately 50 residues) domain from CREB-binding protein (CBP, also called CREBBP), which is a transcriptional coactivator with histone acetylase activity and present in all animals [2]. The NCBD domain has a hydrophobic core, but is very dynamic and even verging on being intrinsically disordered [3–5]. It has therefore been described as "molten-globule-like" [6–9]. The fact that NCBD possesses properties of both intrinsically disordered and globular proteins makes it an interesting system for assessing the role of dynamics in protein structure and how it modulates function. We have previously subjected NCBD to "evolutionary biochemistry" [10,11]. Thus, we resurrected and characterized ancestral versions of NCBD and its protein ligand CBP-Interacting Domain (CID) from the NCOA1, 2 and 3 protein family (with three members in human called Src1, Tif2 and ACTR, respectively) to address the evolution of affinity in this protein-protein interaction. The oldest "maximum likelihood" (ML) NCBD variant we could resurrect was present before the Cambrian period, some 540-600 million years ago, in the common ancestor of all present-day animals with bilaterian symmetry. These animals are divided into deuterostomes and protostomes (D/P) and the ancestral NCBD domain is denoted $NCBD_{D/P}^{ML}$. It was shown to bind the CID domain with relatively low affinity (K_d ~ 1-5 μM) whereas more recent variants of NCBD in the vertebrate lineage

(ca. 440 million years ago) had acquired present-day affinity for CID (K_d ~ 0.1-0.2 μM). Structural characterization of ancestral and present-day CID/NCBD complexes showed that the increase in affinity observed for more modern NCBD variants was due to a combination of factors. These include several new interactions driving structural and dynamic changes in the complex, such as formation of a third α helix in CID upon interaction[11]. In the present paper, we investigate the structural and biophysical properties of the ligand-free state of the ancestral Cambrian-like low-affinity $NCBD_{D/P}^{ML}$ and compare it to that of high-affinity present-day $NCBD_{Human}$ to track the evolution of structure and dynamics and their relation to function. The comparison reveals that the overall properties of the ancestral $NCBD_{D/P}^{ML}$ domain are very similar to those of the present-day $NCBD_{Human}$. In general, conservation implies function. Thus, this conservation of a highly dynamic structure and molten-globule like properties suggest that these features are under positive selection and hence that dynamics are indeed important for the function of NCBD.

Results

The structures of ancient $NCBD_{D/P}$ and extant $NCBD_{Human}$ are similar but not identical. Because of its dynamic properties it is challenging to solve the structure of ligand-free NCBD [6]. First, we performed experiments with ^{15}N labeled $NCBD_{D/P}^{ML}$. We discovered by serendipity that low pH significantly improved the quality of the $^{1}H^{15}N$ heteronuclear single quantum coherence (HSQC) spectrum resulting in well dispersed peaks for $NCBD_{D/P}^{ML}$ (Fig. 1), and allowing for a near complete assignment of the backbone and side-chain residues from NMR 3D experiments using ^{13}C and ^{15}N

labeled samples. The NMR structure was thus determined at pH 2.4 by measuring 3D 1H-1H $^{13}C/^{15}N$ resolved NOESY experiments for distance restraints determination and $^{3}J_{HNHA}$ for dihedral angle determination. The structure of $NCBD_{D/P}^{ML}$ was found to be overall consistent with that of the free[6] and bound $NCBD_{Human}$[4] and bound $NCBD_{D/P}^{ML}$ determined at pH 6.8[11] (Fig. 1). However, a detailed comparison between $NCBD_{D/P}^{ML}$ and $NCBD_{Human}$ shows that the orientation of helix 1 (α1) and helix 3 (α3) differ between the structures. Upon structural alignment of $NCBD_{Human}$ and $NCBD_{D/P}^{ML}$, we observed that both the α1 and α3 from $NCBD_{D/P}^{ML}$ are slightly tilted away relative to that of $NCBD_{Human}$. We also observed that for CID-bound $NCBD_{D/P}^{ML}$ and free $NCBD_{D/P}^{ML}$ the α3 displayed structural rearrangements (Fig. 2).

The amino acid sequence of $NCBD_{D/P}^{ML}$ was reconstructed using phylogenetic methods[4]. The probability is low that the resulting maximum likelihood sequence is identical to the actual sequence present in the ancestor. Thus, $NCBD_{D/P}^{ML}$ is rather one of a large number of likely ancestral variants with similar properties. We wondered if errors in the sequence of $NCBD_{D/P}^{ML}$ resulted in the differences we observe in helix α1 and α3 compared to $NCBD_{Human}$. To investigate this further, and to test how structurally robust the dynamic NCBD domain is, we performed a control experiment where we expressed, purified and characterized an alternative $NCBD_{D/P}$ variant denoted $NCBD_{D/P}^{AltAll}$. In this $NCBD_{D/P}^{AltAll}$ variant, all residues with a posterior probability lower than around 0.9 [10] were replaced with the second most likely residue at that position. For example, residue 2107 is His with 87% probability and Gln with 12% probability in the ancestral $NCBD_{D/P}$. Thus, $NCBD_{D/P}^{ML}$ has a His in position 2107 and $NCBD_{D/P}^{AltAll}$ a Gln residue. $NCBD_{D/P}^{AltAll}$ can be considered a "worst case scenario" variant and a good control of the robustness of conclusions drawn from resurrection

experiments. If ML and AltAll variants display similar properties it is very likely that the actual ancestral protein shares these properties as well. Because of this, an AltAll variant is a convenient alternative to combinations of point mutations [12]. $NCBD_{D/P}^{AltAll}$ contains six substitutions as compared to $NCBD_{D/P}^{ML}$ (Fig. 1). Three of the differences between $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$ are in the N-terminus, but only one of these is in a structured region, a Gln2065→Pro at the beginning of α1. The other three differences are Ser2078→Asn in the loop between α1 and α2, a solvent exposed Gln2088→His in α2 and His2108→Gln at the C-terminus. We solved the NMR structure of $NCBD_{D/P}^{AltAll}$ (Fig. S1) and found that it folds into a similar structure as the other NCBD domains. The overall RMSD between $NCBD_{D/P}^{AltAll}$ and $NCBD_{D/P}^{ML}$ was 7.9 Å. As a comparison, the overall RMSD between $NCBD_{Human}$ and $NCBD_{D/P}^{AltAll}$ or $NCBD_{D/P}^{ML}$ were 4.5 Å and 7.0 Å, respectively. Detailed inspection of the structures revealed differences in the backbone between $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$, especially in the orientations of the helices. An overlay of the structures of $NCBD_{D/P}^{AltAll}$, $NCBD_{D/P}^{ML}$, and $NCBD_{Human,}$ shows differences between α1, α2 and α3 (Fig. S1). One reason for the differences is that there are several long-range NOEs in $NCBD_{D/P}^{AltAll}$ that are not present in $NCBD_{D/P}^{ML}$, between the following pairs of residues: 2068/2082, 2069/2080, 2070/2083 and 2071/2099. There is also an NOE between Ile2062 Hγ, and Pro2065 Hα, two of the substituted residues in $NCBD_{D/P}^{AltAll}$. However, there appears to be a consistency in the hydrophobic core of both $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$. The single Phe residue at position 2100 is packed and buried in a similar fashion in both $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$. Phe2100 is stabilized by Leu2090 in $NCBD_{D/P}^{ML}$, and both Ile2089 and Leu2090 in $NCBD_{D/P}^{AltAll}$ (Fig. 3). Interestingly, this is the same kind of packing that was observed in the free $NCBD_{human}$ domain [6], suggesting conserved features in the hydrophobic core.

Despite the differences discussed in the previous paragraph, the overall biophysical properties of $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$ were found to be similar (Fig. S2) corroborating the overall similarity seen with the NMR structures. Specifically, to test the affinity of $NCBD_{D/P}^{AltAll}$ for CID, we expressed and purified ML and AltAll versions of its ligand CID_{1R} based on our previous reconstruction[6]. K_d values of all four combinations ($NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$ with CID_{1R}^{ML} and CID_{1R}^{AltAll}) were between 2-9 μM (Fig. S2a), which is in the same range as previous estimates of affinity of the ancestral complex [10,11]. Furthermore, the stability as monitored by urea denaturation (Fig. S2b) and thermal denaturation (Fig. S2c-d) were similar for $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$. Thus, based on the highly similar biophysical properties of $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$ we conclude that our conclusions are robust with regard to the uncertainty in the reconstructed sequence.

Ancient $NCBD_{D/P}$ and extant $NCBD_{Human}$ display similar pH dependence with regard to structure and stability. Since the $NCBD_{D/P}^{ML}$ structure was determined at low pH, we next investigated how pH influences the fold and stability. First, we performed ^{1}H-^{15}N HSQC-monitored pH titrations of $NCBD_{D/P}^{ML}$ and compared them to those of $NCBD_{Human}$ at different pH values (Fig. 4). The NMR titration experiments showed that $NCBD_{D/P}^{ML}$ is as folded at low pH as it is at high pH values. However, the peaks (^{1}H-^{15}N) are much sharper (reduced line-width) at lower pH values indicating less conformational heterogeneity. Second, the CD-monitored urea denaturation experiments showed identical stability pattern at two pH values (3.0 and 7.4) for both ancient $NCBD_{D/P}^{ML}$ and extant $NCBD_{Human}$ (Fig. 4c and d, Supplementary table S1). This indicates that while at lower pH the $NCBD_{D/P}^{ML}$ still exhibits the same stability as

at high pH values, there is a shift in conformational heterogeneity towards a single state at low pH. Furthermore, whereas the apparent thermodynamic stability differs by approximately 1 kcal mol^{-1} for ancient NCBD$_{D/P}^{ML}$ and extant NCBD$_{Human}$, both are largely unaffected by a drop in pH to 3.0 (Supplementary Table S1). The observed difference in stability between NCBD$_{D/P}^{ML}$ and NCBD$_{Human}$ appears to be a robust result. Previous experiments suggested that the ancestral NCBD$_{D/P}^{ML}$ was slightly less stable than younger and extant NCBD variants as judged by urea denaturation experiments at pH 7.5 [10]. However, the broad unfolding transition of NCBD resulting from its small hydrophobic core leads to a large error in the $m_{D\text{-}N}$ value, which makes it difficult to unequivocally determine the thermodynamic stability of the domain. However, all present results (including experiments on the stabilized NCBD$_{D/P}^{ML}$ variant NCBD$_{D/P}^{T2073W}$ (Supplementary Table S1) agree well with the previous experiments and we can therefore conclude with greater confidence that the ancestral NCBD$_{D/P}^{ML}$ was less thermodynamically stable than evolutionarily younger variants. Interestingly, with some exceptions [13,14], ancestral reconstructions often yield proteins with higher thermostability, which correlates with thermodynamic stability, than extant proteins, either due to a bias for stability in the reconstruction or due to higher temperatures in past times [15–19], like during Cambrian [20,21]. In any case, ancestral NCBD appears to have populated the native state to a lesser extent than present day human NCBD.

Ancient NCBD$_{D/P}^{ML}$ and extant NCBD$_{Human}$ display similar temperature-dependent structural changes. Previous experiments demonstrated that NCBD$_{Human}$ shows an apparent non-cooperative unfolding behavior when subjected to increasing temperatures, as monitored by CD [6,22]. This behavior may result from a low $\Delta H_{D\text{-}N}$

value rather than a true non-cooperative unfolding. (The urea denaturation experiments are consistent with a cooperative two-state unfolding.) We compared ^{1}H-^{15}N HSQC and CD spectra, respectively, of $NCBD_{D/P}^{ML}$ and $NCBD_{Human}$ at different temperatures (Fig. 5). In agreement with the previous data we observed that both variants unfold non-cooperatively with temperature as monitored by the increase in molar ellipticity at 222 nm suggesting loss of α-helix as the temperature is increased. However, for both $NCBD_{D/P}^{ML}$ and $NCBD_{Human}$ the signal is not completely lost even at 363 K (90°C) underscoring that this is not a typical native to highly-disordered-denatured state transition. Furthermore, ^{1}H-^{15}N HSQC spectra of $NCBD_{D/P}^{ML}$ at temperatures up to 333 K (60°C) show well dispersed peaks indicating a presumably globular, collapsed architecture (Fig. S3) and corroborating the notion that NCBD retains structure at high temperatures.

Rate constants for (un)folding of ancient $NCBD_{D/P}^{ML}$ and extant $NCBD_{Human}$. The folding of human NCBD is complex, and, based on kinetic experiments on a human NCBD variant, we proposed a three-state system involving 'native' NCBD (*i.e.*, the solved structure), an intermediate state with unknown structure and a denatured state, which likely retains significant structure under native conditions [23]. Rate constants in the same order was determined by relaxation dispersion NMR experiments by Kjaergaard et al.[8] It is clear from the NMR experiments that low pH promotes the native state of NCBD and that pH does not affect the thermodynamic stability as monitored by urea denaturation experiments to any significant extent, neither for $NCBD_{D/P}^{ML}$ nor $NCBD_{Human}$. Previous folding experiments on $NCBD_{Human}$ [23] involving jumps from low to neutral pH thus resulted in small shifts in the equilibria between the native, intermediate and denatured states yielding the observed kinetic transients. In fact, a

likely explanation for the better NMR spectra at low pH is that there is less of the intermediate state under these conditions.

To further compare the biophysical properties of ancient $NCBD_{D/P}^{ML}$ and extant $NCBD_{Human}$ we performed kinetic folding experiments using stopped-flow as well as temperature-jump fluorescence spectroscopy. To achieve this we introduced a Trp at position 2073 in $NCBD_{D/P}^{ML}$ as previously done in the folding studies of $NCBD_{Human}$ [23]. This modification resulted in a thermodynamic stabilization of around 1 kcal mol^{-1} for both $NCBD_{D/P}^{T2073W}$ and $NCBD_{Human}^{T2073W}$ (Fig. 6a-b, Supplementary Table S1). Previous folding experiments were performed on a slightly longer $NCBD_{Human}$ variant (residues 2058-2116), which was used in the original studies of CID/NCBD [22]. In our resurrection studies, we used a shorter version of $NCBD_{Human}$, only containing the evolutionarily conserved region (residues 2062-2109). Therefore, we first repeated the temperature-jump folding experiments for $NCBD_{Human}^{T2073W}$ and showed that the shorter construct used in our evolutionary studies displayed similar kinetics as the longer construct (Fig. 6c). We then performed folding experiments with $NCBD_{D/P}^{T2073W}$. The kinetic transients of $NCBD_{D/P}^{T2073W}$ from temperature jump experiments had lower signal-to-noise than those of $NCBD_{Human}^{T2073W}$, but observed rate constants were in the same range (Fig. 6d). To investigate the folding kinetics further we used stopped-flow fluorimetry, which has a lower time resolution but can be used to compare folding rate constants for $NCBD_{D/P}^{T2073W}$ and $NCBD_{Human}^{T2073W}$ at low temperature (Fig. 6e-f). The stopped-flow folding experiments were conducted by making pH jumps from a 5 mM HCl solution (pH ~ 2) to a buffer solution of 20 mM sodium phosphate (pH 7.4, 150 mM NaCl, 1 M TMAO) and monitoring the relaxation through a 330 nm band-pass filter at 4°C. The low temperature is necessary to reduce

the observed rate constant such that the folding or other conformational transitions take place in the millisecond time window. We observed two kinetic phases for $NCBD_{Human}^{T2073W}$ (one fast phase with a negative amplitude and a second slower phase with a positive amplitude) but only one kinetic phase for $NCBD_{D/P}^{T2073W}$ (with a positive amplitude). The values of the observed rate constants for $NCBD_{Human}^{T2073W}$ (~300 s^{-1} and ~70 s^{-1}, respectively) were similar to those observed previously for the longer construct [23]. The single value obtained for the $NCBD_{D/P}^{T2073W}$ was ~140 s^{-1}. Trp fluorescence is a very sensitive but crude structural probe. Thus, the number of kinetic phases in folding experiments could be dependent on small structural rearrangements around the Trp. Nevertheless, the observed rate constant for $NCBD_{D/P}^{T2073W}$ in the stopped flow pH jump experiment is in the same range as the two rate constants observed under the same conditions for human NCBD.

Discussion

The biophysics and molecular evolution of NCBD is interesting for several reasons, which all relate to the dynamic properties of NCBD. Firstly, while it has a hydrophobic core and globular shape, NCBD is a very dynamic protein domain with characteristics of an IDP [3,6–9,22,24]. Secondly, NCBD has several binding partners including the three NCOA transcriptional co-regulator paralogs [3,25], the transcription factors p53 [26], Ets-2 [27] and interferon factor-3 [28], and viral proteins such as Tax [29] from T-cell leukemia virus. Intriguingly, NCBD displays great conformational plasticity as shown by its complex with interferon factor-3 where the helices of NCBD adopt a distinct conformation as compared to the complex with CID [30]. Thirdly, evolutionary snapshots

of the complex between CID and NCBD suggest that several new intermolecular contacts and conformational plasticity increased the affinity when going from the Cambrian-like $NCBD_{D/P}^{ML}$ (K_d ~ 1-5 μM) to younger variants (K_d ~ 0.2 μM) [10,11,31]. Fourthly, while affinity and structural order of modern NCBD/CID complexes were increased, the transition state of the folding-induced binding displayed more conformational heterogeneity as compared to the ancestral complex [32]. Our present results, how the free, unbound state of NCBD has evolved during the transition from a low-affinity to a high-affinity CID binder relate to all four points.

The most striking finding is that $NCBD_{D/P}^{ML}$ displays very similar molten-globule like properties as the previously characterized human NCBD [6,8]. This conservation of properties, borderline between a globular domain and an IDP, suggests that the dynamics of NCBD is under positive selection and therefore of functional importance. In general, structural plasticity is regarded as one main functional feature of IDPs allowing for multiple protein-protein interactions [33]. NCBD interacts with several protein domains, both folded and disordered ones, and the dynamic nature of NCBD likely increases its propensity to adapt to different partners. Moreover, NCBD interactions result in coupled folding and binding. Many such reactions have recently been shown to occur via a mechanism denoted templated folding, where the binding partner influences the folding pathway [34–36]. Indeed, this was observed for the interaction between human NCBD and CID [32,37,38]. Thus, our present findings suggest that templated folding may be a general feature of binding reactions with NCBD where the conserved dynamics facilitate sampling of free states, heterogeneity of bound states, as well as the pathway between free and bound states.

Going into details of the NMR model, we note that while overall structure is robust to errors in the predicted sequence as shown by the comparison of the NMR models of free $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$, there are differences in for example helix orientations. Globular domains with a sequence identity as high as these NCBD variants (77-88%) would typically have identical structures due to a highly funneled energy landscape [39,40], while proteins with larger structural heterogeneity have less funneled landscapes [1]. Indeed, the differences between the structures of $NCBD_{Human}$, $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$, in particular regarding the angles of the helices in relation to each other, are consistent with a less funneled energy landscape resulting in conformational heterogeneity [9]. This heterogeneity is also consistent with the conserved dynamic properties of NCBD. In conclusion, our data suggest evolutionarily conserved dynamics of NCBD that underlie functional plasticity in ground state conformations, templated binding-and-folding and a propensity to interact with several binding partners.

Materials and methods

Protein sequences, expression and purification. The reconstruction of ancestral sequences was previously published [10]. The AltAll versions of $NCBD_{D/P}$ (Fig. 1) and CID_{1R} used in the present paper was based on posterior probabilities from the reconstruction with a cutoff of 91%. One position in $NCBD_{D/P}$ (Thr2062) is very uncertain. A Thr residue (15%) was originally chosen at this position in $NCBD_{D/P}^{ML}$ to avoid a hydrophobic residue, although Ile (19%) has a higher posterior probability. In the current $NCBD_{D/P}^{AltAll}$ residue 2062 is Ile. NCBD and CID variants were expressed

in *E. coli* from a pRSETA plasmid and purified as previously described [32] Concentrations were determined by absorbance at 280 nm for NCBD (using calculated extinction coefficients based on amino acid composition) or 205 nm for CID using an extinction coefficient of 250,000 $M^{-1}cm^{-1}$. Purity was checked by SDS-PAGE and identity of the purified proteins was confirmed by MALDI-TOF mass spectrometry.

Circular dichroism. Far-UV CD experiments were performed using a JASCO J-1500 CD spectrophotometer with a Peltier temperature control system. The protein concentration in all CD experiments was 18-20 μM. The buffer solution was either 20 mM sodium phosphate (pH 7.4), 150 mM NaCl or 50 mM potassium formate (pH 3.0), 150 mM NaCl, and CD spectra were recorded at 277 K. The scanning speed was 50 nm/min, bandpass 1 nm and integration time 1 second. Each reported spectrum is an average of 3-4 recorded spectra. In the equilibrium thermal denaturation experiments the proteins were dissolved in 50 mM potassium formate buffer (pH 3.0), 150 mM NaCl and the sample was heated from 277 to 368 K. The heating rate was 1 K/min, waiting time 5 seconds and the denaturation was monitored at 222 nm. During the heating, spectra (average of three measurements) were taken at 283, 303, 323, 343 and 363 K, respectively. In the equilibrium urea denaturation experiments the proteins were denatured in 0-6.6 M urea at 277 K and the unfolding was monitored at 222 nm. The data was analyzed in GraphPad Prism and fitted by non-linear regression to a two-state equilibrium model [41]. Because of the broad unfolding transition and lack of native baseline for several variants, four parameters were shared among all urea denaturation data sets in the curve fitting (Fig. S4, Supplementary Table 1). One of these parameters was the molar ellipticity of the native state, with the reasonable assumption that the secondary structures of the native states are similar for all variants. The other

parameters were the slopes of the denatured and native baselines as well as the $m_{D\text{-}N}$ value (assuming a similar change in solvent accessible surface area upon denaturation for all variants).

Temperature jump experiments. Temperature jump experiments were conducted using 100-200 µM protein on a TJ-64 temperature jump system (TgK Scientific). The buffer solution was 20 mM sodium phosphate (pH 7.4), 150 mM NaCl. The proteins were subjected to temperature jumps of 2-8.5 K to different target temperatures. The data were fitted in the Kinetic Studio software (TgK scientific) to a double exponential function, to account both for the fast phase (relaxation time ~ 80 µs), which corresponds to the heating time of the instrument, and to the slow phase, which corresponds to the conformational transition in the protein. A Trp variant of $NCBD_{D/P}^{ML}$ ($NCBD_{D/P}^{T2073W}$) was chosen for kinetic experiments since a Trp in this position was previously used for human NCBD. We engineered a Trp at other positions in $NCBD_{D/P}^{ML}$ and obtained rate constants in the same range for all variants (Supplementary Table S2).

Stopped flow experiments. The stopped flow experiments were performed using an upgraded SX-17MV stopped flow spectrophotometer (Applied Photophysics) and the measurements were conducted at 277 K. Proteins were dissolved in 5 mM HCl (pH ~ 2) and rapidly mixed 1:1 with buffer solutions complemented with trimethylamine N-oxide (TMAO) to a final concentration of 20 mM sodium phosphate (pH 7.4), 150 mM NaCl, 1 M TMAO. The final protein concentration was 5 µM. The excitation wavelength was 280 nm and the resulting fluorescence emission was detected after passage through a 330 nm bandpass filter. The transients were analyzed in GraphPad Prism and fitted to a single or double exponential function.

NMR experiments. NMR experiments were acquired on Bruker 600, 700 and 900 MHz spectrometers equipped with triple resonance cryogenic temperature probes at 298 K except otherwise stated. The final NMR samples contained 500 µM protein to which 0.01% NaN_3 and 5% D_2O was added. Experiments for assignment and subsequent structure determination were done at pH 2.4 (HCl adjusted with NaOH) for $NCBD_{D/P}^{ML}$ and pH 6.8 for $NCBD_{D/P}^{AltAll}$. The following NMR experiments were recorded for assignment and subsequent structure determination: standard 2D ^{1}H-^{15}N HSQC, 3D HNCACB ^{15}N-resolved [^{1}H-^{1}H]-NOESY-HSQC, ^{15}N-TOCSY-HSQC ^{13}C-resolved [^{1}H-^{1}H]-NOESY-HSQC, HCCH-TOCSY and ^{13}C-resolved [^{1}H-^{1}H]-TOCSY-HSQC. Phi-angle restraining $^{3}J_{HNHA}$ couplings were determined from 3D HNHA type experiments using quantitative -*J* intensity modulated experiments [42]. The temperature variation experiments were performed in 20 mM sodium formate, pH 3.0. All experiments were processed with nmrpipe [43] and analysed with CCPnmr [44].

Structure calculation. Structure calculations were done using the CYANA 3.97 [45] package as follows: Initially, cross peaks were converted into upper distance restraints following an automated process in CYANA. These distance restraints together with φ/ψ dihedral angles determined from C^{α}-chemical shifts and $^{3}J_{HNH}\alpha$ (measured) were used as input for the initial structure calculations. The structures were calculated with 200,000 torsion angle dynamics steps for 100 conformers starting from random torsion angles by simulated annealing. 20 conformers with the lowest target function values were selected and analyzed. The structural statistics together with all input data for the structure calculations are presented in Supplementary Table S3. Assignments and

structural coordinates have been deposited in the RCSB protein data bank with pdb code: 7OSR ($NCBD_{D/P}^{ML}$) and 7OSW ($NCBD_{D/P}^{AltAll}$).

Acknowledgements

This work was supported by the Wenner-Gren foundation WG-17 returning grants to C. N. C., and by the Swedish Research Council (2020-04395 to P.J.) and Knut and Alice Wallenberg foundation (Evolution of new genes and proteins, to P.J.). We used the NMR Uppsala infrastructure, which is funded by the Department of Chemistry - BMC and the Disciplinary Domain of Medicine and Pharmacy. We thank Magnus Kjaergaard for valuable comments on the NMR experiments.

References

1. Gianni, S. *et al.* Fuzziness and Frustration in the Energy Landscape of Protein Folding, Function, and Assembly. *Acc Chem Res* **54**, 1251–1259 (2021).
2. Dyson, H. J. & Wright, P. E. Role of Intrinsic Protein Disorder in the Function and Interactions of the Transcriptional Coactivators CREB-binding Protein (CBP) and p300. *J. Biol. Chem.* **291**, 6714–22 (2016).
3. Lin, C. H. *et al.* A small domain of CBP/p300 binds diverse proteins: solution structure and functional studies. *Mol Cell* **8**, 581–590 (2001).
4. Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).
5. Ebert, M.-O., Bae, S.-H., Dyson, H. J. & Wright, P. E. NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. *Biochemistry* **47**, 1299–308 (2008).
6. Kjaergaard, M., Teilum, K. & Poulsen, F. M. Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12535–12540 (2010).
7. Kjaergaard, M., Poulsen, F. M. & Teilum, K. Is a malleable protein necessarily highly dynamic? The hydrophobic core of the nuclear coactivator binding domain is well ordered. *Biophys J* **102**, 1627–1635 (2012).
8. Kjaergaard, M., Andersen, L., Nielsen, L. D. & Teilum, K. A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition. *Biochemistry* **52**, 1686–1693 (2013).

9. Papaleo, E., Camilloni, C., Teilum, K., Vendruscolo, M. & Lindorff-Larsen, K. Molecular dynamics ensemble refinement of the heterogeneous native state of NCBD using chemical shifts and NOEs. *PeerJ* **6**, e5125 (2018).
10. Hultqvist, G. *et al.* Emergence and evolution of an interaction between intrinsically disordered proteins. *Elife* **6**, e16059 (2017).
11. Jemth, P. *et al.* Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins. *Sci. Adv.* **4**, eaau4130 (2018).
12. Eick, G. N., Bridgham, J. T., Anderson, D. P., Harms, M. J. & Thornton, J. W. Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty. *Molecular Biology and Evolution* msw223 (2016) doi:10.1093/molbev/msw223.
13. Hart, K. M. *et al.* Thermodynamic system drift in protein evolution. *PLoS Biol* **12**, e1001994 (2014).
14. Laursen, L., Čalyševa, J., Gibson, T. J. & Jemth, P. Divergent Evolution of a Protein-Protein Interaction Revealed through Ancestral Sequence Reconstruction and Resurrection. *Mol Biol Evol* **38**, 152–167 (2021).
15. Williams, P. D., Pollock, D. D., Blackburne, B. P. & Goldstein, R. A. Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput Biol* **2**, e69 (2006).
16. Risso, V. A., Gavira, J. A., Gaucher, E. A. & Sanchez-Ruiz, J. M. Phenotypic comparisons of consensus variants versus laboratory resurrections of Precambrian proteins. *Proteins* **82**, 887–896 (2014).
17. Romero-Romero, M. L., Risso, V. A., Martinez-Rodriguez, S., Ibarra-Molero, B. & Sanchez-Ruiz, J. M. Engineering ancestral protein hyperstability. *Biochem J* **473**, 3611–3620 (2016).

18. Trudeau, D. L., Kaltenbach, M. & Tawfik, D. S. On the Potential Origins of the High Stability of Reconstructed Ancestral Proteins. *Mol Biol Evol* **33**, 2633–2641 (2016).

19. Wheeler, L. C., Lim, S. A., Marqusee, S. & Harms, M. J. The thermostability and specificity of ancient proteins. *Curr Opin Struct Biol* **38**, 37–43 (2016).

20. Hearing, T. W. *et al.* An early Cambrian greenhouse climate. *Sci Adv* **4**, eaar5690 (2018).

21. Wotte, T., Skovsted, C. B., Whitehouse, M. J. & Kouchinsky, A. Isotopic evidence for temperate oceans during the Cambrian Explosion. *Sci Rep* **9**, 6330 (2019).

22. Demarest, S. J. *et al.* Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549–553 (2002).

23. Dogan, J., Toto, A., Andersson, E., Gianni, S. & Jemth, P. Activation Barrier-Limited Folding and Conformational Sampling of a Dynamic Protein Domain. *Biochemistry* **55**, 5289–5295 (2016).

24. Demarest, S. J., Deechongkit, S., Dyson, H. J., Evans, R. M. & Wright, P. E. Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Sci.* **13**, 203–210 (2004).

25. Xu, J. & Li, Q. Review of the in vivo functions of the p160 steroid receptor coactivator family. *Mol. Endocrinol.* **17**, 1681–1692 (2003).

26. Livengood, J. A. *et al.* p53 Transcriptional activity is mediated through the SRC1-interacting domain of CBP/p300. *J Biol Chem* **277**, 9054–9061 (2002).

27. Jayaraman, G. *et al.* p300/cAMP-responsive element-binding protein interactions with ets-1 and ets-2 in the transcriptional activation of the human stromelysin promoter. *J Biol Chem* **274**, 17342–17352 (1999).

28. Matsuda, S. *et al.* A Conserved alpha-helical motif mediates the binding of diverse nuclear proteins to the SRC1 interaction domain of CBP. *J Biol Chem* **279**, 14055–14064 (2004).

29. Scoggin, K. E., Ulloa, A. & Nyborg, J. K. The oncoprotein Tax binds the SRC-1-interacting domain of CBP/p300 to mediate transcriptional activation. *Mol Cell Biol* **21**, 5520–5530 (2001).

30. Qin, B. Y. *et al.* Crystal structure of IRF-3 in complex with CBP. *Structure* **13**, 1269–1277 (2005).

31. Karlsson, E., Lindberg, A., Andersson, E. & Jemth, P. High affinity between CREBBP/p300 and NCOA evolved in vertebrates. *Protein Sci.* **29**, 1687–1691 (2020).

32. Karlsson, E. *et al.* Mapping the transition state for a binding reaction between ancient intrinsically disordered proteins. *J Biol Chem* **295**, 17698–17712 (2020).

33. Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* **16**, 18–29 (2015).

34. Toto, A. *et al.* Molecular Recognition by Templated Folding of an Intrinsically Disordered Protein. *Sci. Rep.* **6**, 21994 (2016).

35. Toto, A. *et al.* Templated folding of intrinsically disordered proteins. *J. Biol. Chem.* **295**, 6586–6593 (2020).

36. Wu, D., Zhou, H.-X., D, W. & HX., Z. Designed Mutations Alter the Binding Pathways of an Intrinsically Disordered Protein. *Sci. Rep.* **9**, 6172 (2019).

37. Jemth, P., Mu, X., Engström, Å. & Dogan, J. A frustrated binding interface for intrinsically disordered proteins. *J. Biol. Chem.* **289**, 5528–5533 (2014).

38. Karlsson, E. *et al.* A structurally heterogeneous transition state underlies coupled binding and folding of disordered proteins. *J. Biol. Chem.* **294**, 1230–1239 (2019).

39. Chothia, C. & Lesk, A. M. The relation between the divergence of sequence and structure in proteins. *EMBO J* **5**, 823–826 (1986).

40. Oliveberg, M. & Wolynes, P. G. The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* **38**, 245–288 (2005).

41. Fersht, A. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding.* (Macmillan, 1999).

42. Lakomek, N.-A., Ying, J. & Bax, A. Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J Biomol NMR* **53**, 209–221 (2012).

43. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293 (1995).

44. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).

45. Güntert, P., Mumenthaler, C. & Wüthrich, K. Torsion angle dynamics for NMR structure calculation with the new program Dyana. *Journal of Molecular Biology* **273**, 283–298 (1997).

Figures

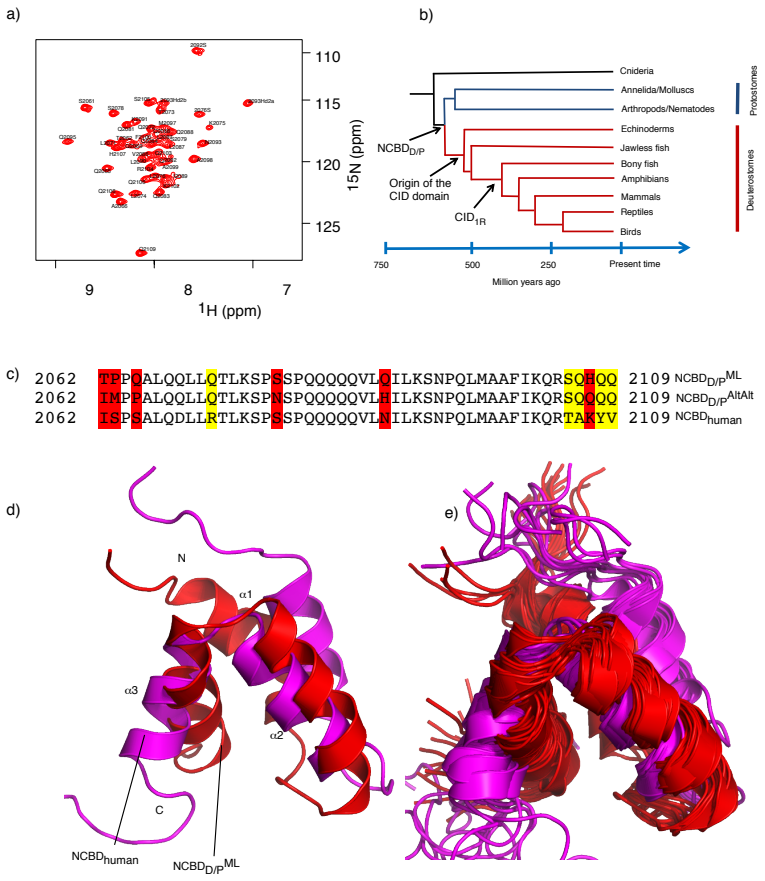

Figure 1. Structures of ancestral and human NCBD domains. (a) ^{1}H-^{15}N HSQC spectrum of $NCBD_{D/P}^{ML}$ at pH 2.4 with assigned residues. All assignments and the structural coordinates have been deposited in the pdb data bank with accession number XXX and pdb code: YYY. [Files are uploaded as supplementary material for review] (b) A schematic tree depicting the relationship between the deuterostome/protostome (D/P) ancestor and modern species. The times on the x-axis are approximate, in particular the divergence of protostomes and deuterostomes. (c) Sequence alignment for $NCBD_{Human}$, $NCBD_{D/P}^{ML}$ $NCBD_{D/P}^{AltAlt}$. The residues differing between $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAlt}$ are marked in red while additional differences to $NCBD_{Human}$ are marked in yellow. (d) Comparison of the NMR structure of ancestral $NCBD_{D/P}^{ML}$ solved at low pH (red), and the previously determined human NCBD

(magenta, pdb code: 2KKJ) [6]. Note that the structure of human NCBD was solved using a slightly longer construct. (e) Overlay of the 20 lowest energy structures of the $NCBD_{D/P}^{ML}$ domain with those of $NCBD_{human}$.

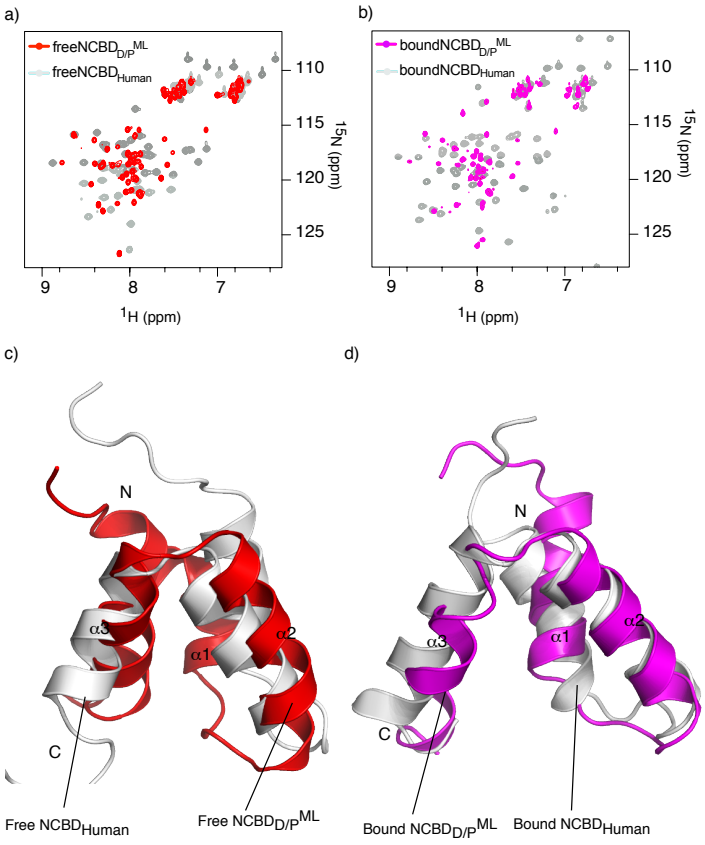

Figure 2. Comparison of the ligand-free and bound ancestral $NCBD_{D/P}^{ML}$ and extant $NCBD_{Human}$. The ligand-bound conformations (pdb codes X and Y) and the structure of free human NCBD (pdb code: 2KKJ) [6] were previously determined. (a) Overlay of the ^{1}H-^{15}N HSQC spectra between free $NCBD_{D/P}^{ML}$ (red) and free $NCBD_{Human}$ (grey). (b) Overlay of the ^{1}H-^{15}N HSQC spectra between CID-bound $NCBD_{D/P}^{ML}$ (magenta) and bound $NCBD_{Human}$ (grey). Structural alignment between (c) free $NCBD_{D/P}^{ML}$ (red) and free $NCBD_{Human}$ (grey), and (d) CID-bound $NCBD_{D/P}^{ML}$ (red) and bound $NCBD_{D/P}^{ML}$ (grey).

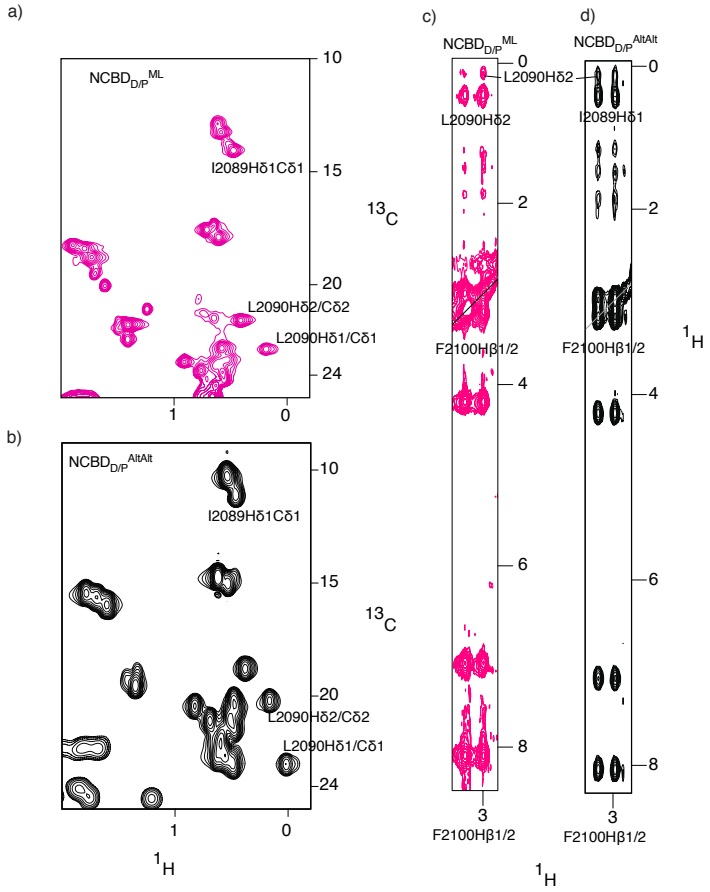

Figure 3. Phe2100 makes NOEs to Ile2089 and Leu2090 in NCBD$_{D/P}^{ML}$ and NCBD$_{D/P}^{AltAlt}$. (a) ^{1}H-^{13}C HSQC spectra of the methyl region of NCBD$_{D/P}^{ML}$, showing the Hδ1-Cδ1 shifts of L2090 and I2089. (b) ^{1}H-^{13}C HSQC spectra of the methyl region of NCBD$_{D/P}^{AltAlt}$, showing the Hδ1-Cδ1 shifts of L2090 and I2089. (c) Strip from a ^{1}H-^{1}H ^{13}C resolved NOESY showing NOEs from the Hβ1/2 atoms of Phe2100 and the associated NOEs to Leu2090Hδ1 for NCBD$_{D/P}^{ML}$. (d) Strip from a ^{1}H-^{1}H ^{13}C resolved NOESY showing NOEs from the Hβ1/2 atoms of Phe2100 and the associated NOEs to Leu2090Hδ1 and Ile2089Hδ1 for NCBD$_{D/P}^{AltAlt}$. These signature NOEs implies that both proteins form a hydrophobic core similar to that of NCBD$_{Human}$.

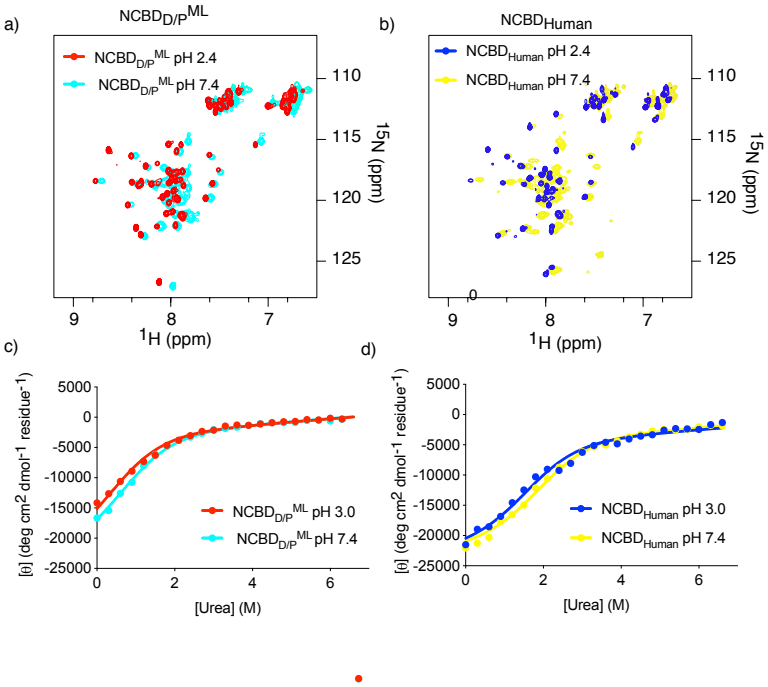

Figure 4. Comparison of ancestral NCBD$_{D/P}^{ML}$ and extant NCBD$_{Human}$ at two different pH values. a) Overlay of the ^{1}H-^{15}N HSQC spectra of NCBD$_{D/P}^{ML}$ at pH 2.4 (red) and at pH 6.8 (cyan). b) Overlay of the ^{1}H-^{15}N HSQC spectra of NCBD$_{Human}$ at pH 2.4 (blue) and at pH 6.8 (yellow). c) Urea denaturation experiments for NCBD$_{D/P}^{ML}$ at pH 3.0 (red) and at pH 7.4 (cyan). d) Urea denaturation experiments for NCBD$_{Human}$ at pH 3.0 (yellow) and at pH 7.4 (blue).

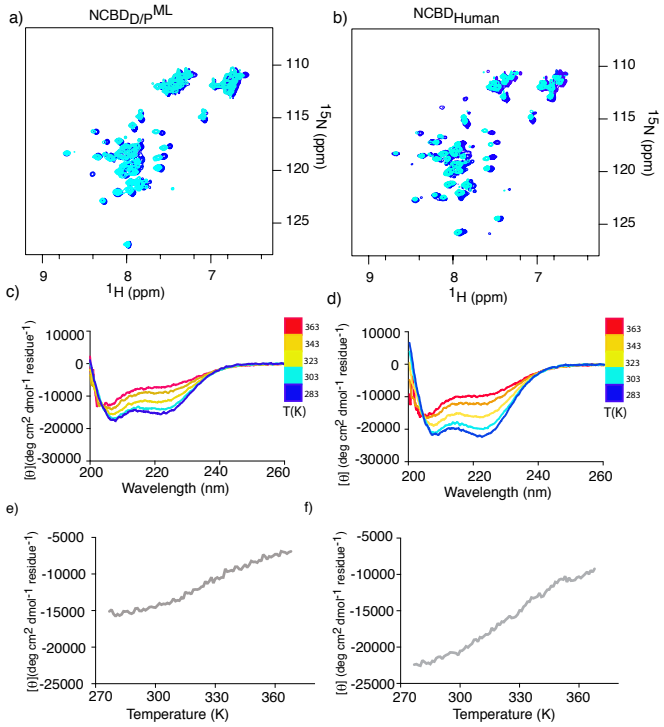

Figure 5. Temperature dependence for ancestral NCBD$_{D/P}^{ML}$ and extant human NCBD. a) Overlay of ^{1}H-^{15}N HSQC spectra of NCBD$_{D/P}^{ML}$ at 283 K (blue) and at 303 K (cyan). b) Overlay of ^{1}H-^{15}N HSQC spectra of NCBD$_{Human}$ at 283 K (blue) and at 303 K (cyan). Circular dichroism spectra at a range of temperatures (283-363 K) for (c) NCBD$_{D/P}^{ML}$ and (d) NCBD$_{Human}$. Molecular ellipticity at 222 nm (probing α-helix content) as a function of temperature for (e) NCBD$_{D/P}^{ML}$ and (f) NCBD$_{Human}$.

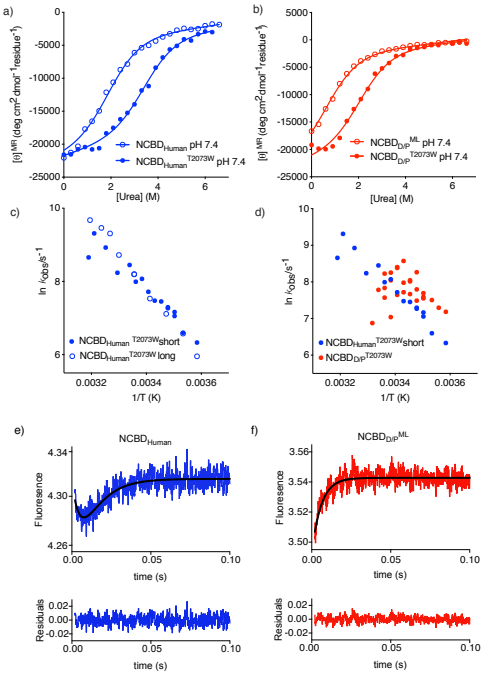

Figure 6. Stability and temperature dependence of kinetic rate constants for NCBD$_{D/P}^{ML}$ and extant human NCBD. (a) Stability measurement of NCBD$_{Human}^{WT}$ (open circles) and NCBD$_{Human}^{T2073W}$ (filled circles) measured by urea denaturation. (b) Stability measurement of NCBD$_{D/P}^{ML}$ (open circles) and NCBD$_{D/P}^{T2073W}$ (filled circles). Parameters from the curve fitting are shown in Supplementary Table S1. (c) Temperature jump kinetic experiments for a longer human NCBD construct used in previous studies [23] (open circles) and the shorter human NCBD variant used in this study (filled circles). Both variants contained the T2073W mutation. (d) A comparison of observed rate constants from temperature-jump experiments for NCBD$_{Human}^{T2073W}$ (blue) and NCBD$_{D/P}^{T2073W}$ (red). All experiments were performed in 20 mM sodium phosphate pH 7.4, 150 mM NaCl. (e-f) Kinetic (un)folding of NCBD$_{Human}^{T2073W}$ and NCBD$_{D/P}^{T2073W}$ measured by stopped flow fluorescence spectroscopy. (e) (Un)folding kinetic transient of NCBD$_{Human}^{T2073W}$ (blue) fitted to a double exponential function (black line). Below each transient are the residuals from the respective fits. (f) Stopped flow (un)folding kinetic transient of NCBD$_{D/P}^{T2073W}$

(red) fitted to a single exponential function (black line). The experiments were performed by jumping from 5 mM HCl (pH ~ 2) to a final buffer concentration of 20 mM sodium phosphate (pH 7.4), 150 mM NaCl, 1 M TMAO. The relaxation was monitored using a 330 band-pass emission filter. Fitting a double exponential function resulted in k_{obs} values of 290 ± 30 and 67 ± 4 s^{-1} for $NCBD_{Human}^{T2073W}$. Fitting a single exponential function yielded a k_{obs} value of 139 ± 7 s^{-1} for $NCBD_{D/P}^{T2073W}$.

Supplementary information

The dynamic properties of a nuclear coactivator binding domain are evolutionarily conserved

Elin Karlsson[1], Frieda A. Sorgenfrei[1,†], Eva Andersson[1], Jakob Dogan[1], Per Jemth[1,*], and Celestine N. Chi[1,2,*]

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden.

[2]Department of Pharmaceutical Biosciences, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden

[†]Present address: acib GmbH, Krenngasse 37, 8010 Graz c/o University of Graz, Institute of Chemistry, NAWI Graz, BioTechMed Graz, Heinrichstrasse 28, 8010 Graz, Austria

*Correspondence to

Celestine Chi, Celestine.Chi@imbim.uu.se

Per Jemth, Per.Jemth@imbim.uu.se

Supplementary figures

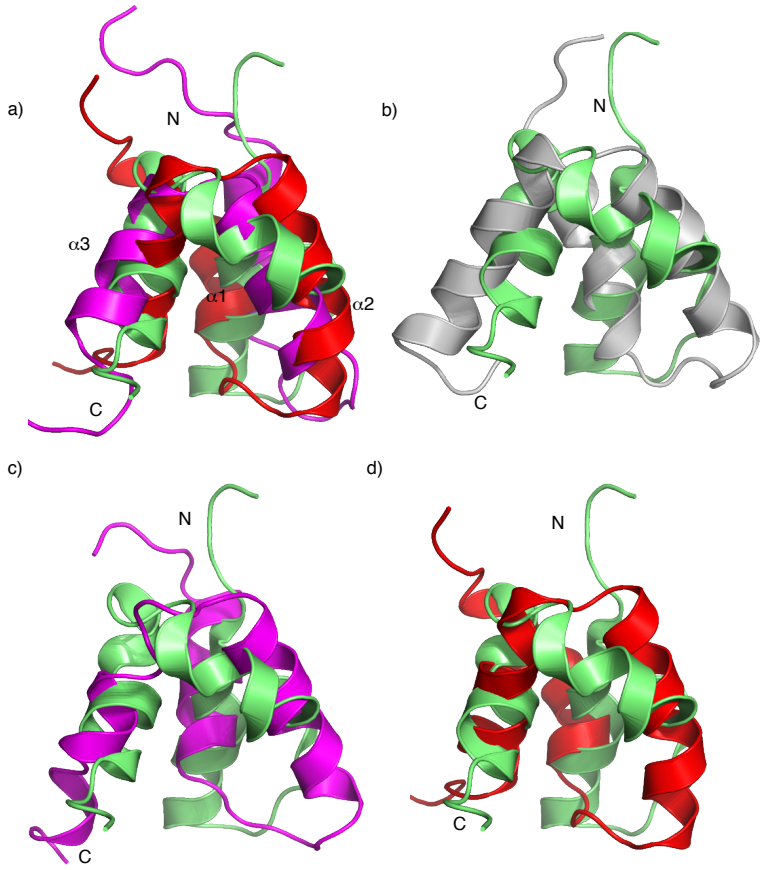

Figure S1. Overlay of the structures of $NCBD_{D/P}^{ML}$, $NCBD_{D/P}^{AltAlt}$ and $NCBD_{Human}$. (a) An overlay of the structures determined from NMR for $NCBD_{D/P}^{ML}$(red), $NCBD_{D/P}^{AltAlt}$ (green), and the previously determined $NCBD_{Human}$ (magenta). (b) Overlay of $NCBD_{D/P}^{AltAlt}$ (green), and the previously determined CID-bound $NCBD_{D/P}^{ML}$ (grey). (c) Overlay of $NCBD_{D/P}^{AltAlt}$ (green), and the previously determined $NCBD_{Human}$ (magenta). (d) Overlay of $NCBD_{D/P}^{AltAlt}$(green) and $NCBD_{D/P}^{ML}$ (red).

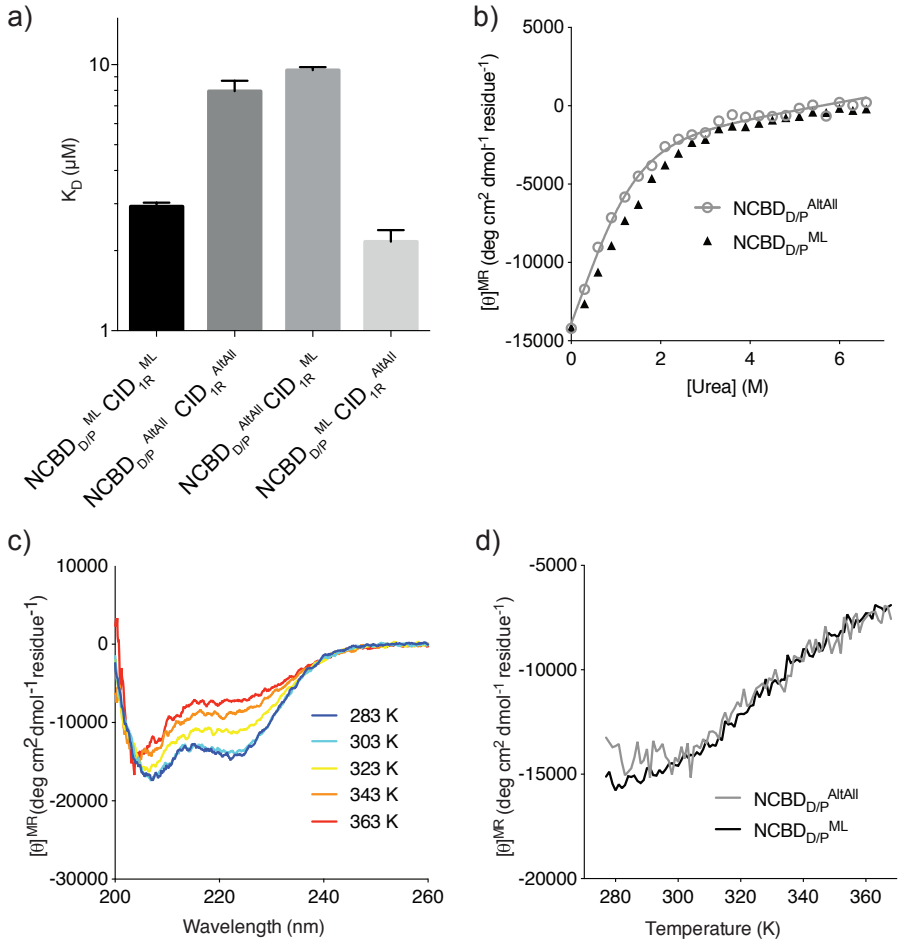

Figure S2. The biophysical properties of $NCBD_{D/P}^{ML}$ are robust to errors in the sequence.

a) Binding affinities of $NCBD_{D/P}^{ML}$ or $NCBD_{D/P}^{AltAlt}$ and CID_{1R}^{ML} or CID_{1R}^{AltAll} measured with ITC. The buffer solution was 20 mM sodium phosphate pH 7.4, 150 mM NaCl. (b) Stability of $NCBD_{D/P}^{AltAll}$ measured by urea denaturation in 50 mM formate (pH 3.0), 150 mM NaCl. The CD signal at 222 nm was monitored and the data was fitted to a two-state function. A reliable $[urea]_{50\%}$ value could not be obtained, although qualitative comparison with $NCBD_{D/P}^{ML}$ indicates that the stability of the two variants is similar. (c-d) Temperature stability for $NCBD_{D/P}^{AltAll}$. The CD signal at 222 nm was monitored and spectra were taken at temperatures 283-363 K.

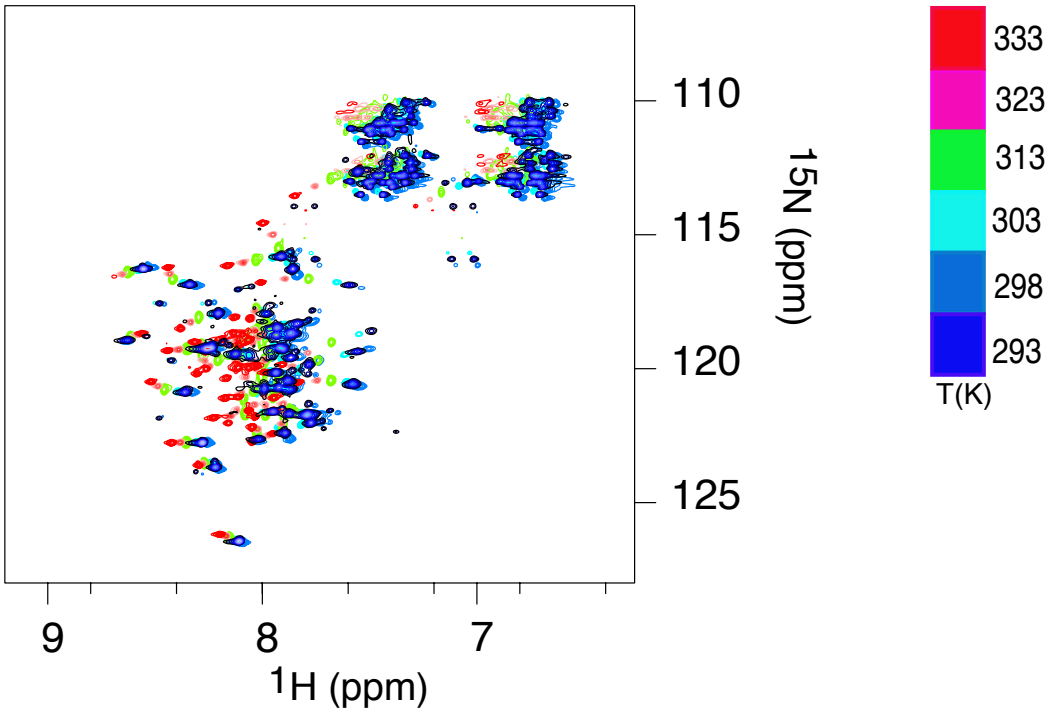

Figure S3. ^{15}N-^{1}H HSQC spectra of NCBD$_{D/P}^{ML}$ at different temperatures. The experiments were performed from 293 K (blue) to 333 K (red).

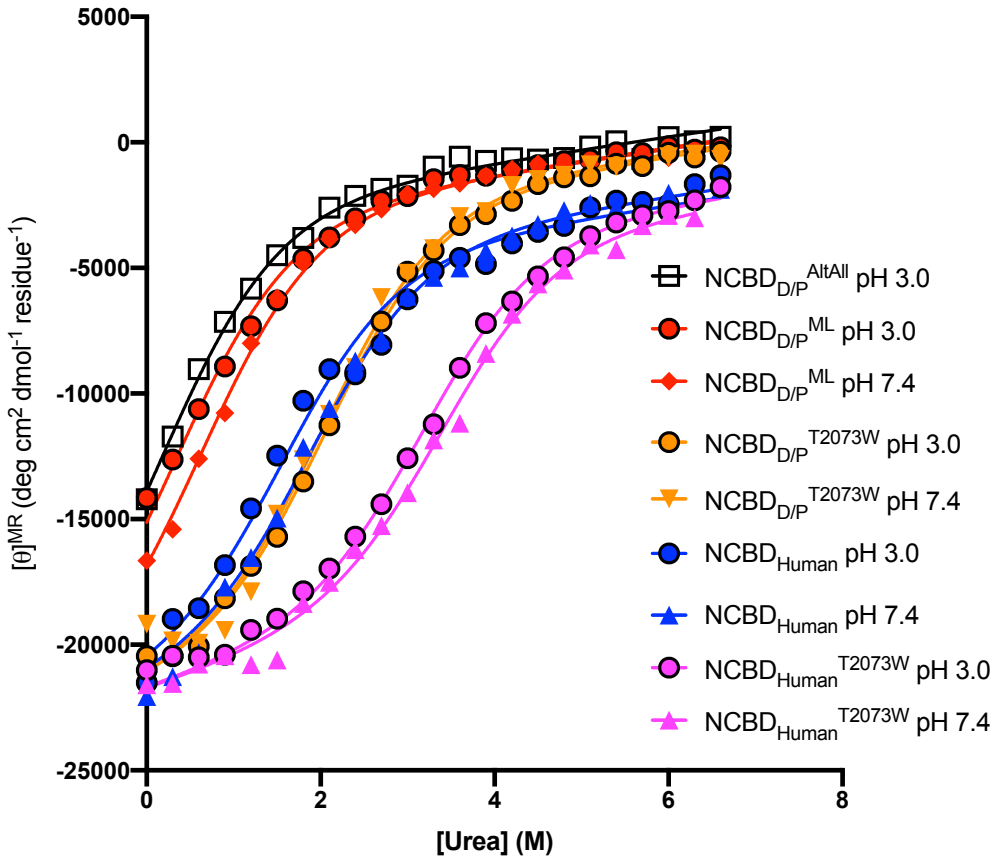

Figure S4. Equilibrium urea denaturation experiments. All data sets were fitted simultaneously to a two-state model in GraphPad Prism with shared parameters for the native state molar ellipticity, the denatured and native baseline slopes and the $m_{D\text{-}N}$ value. Experiments were performed in either 50 mM potassium formate buffer (pH 3.0), 150 mM NaCl or 20 mM sodium phosphate buffer (pH 7.4), 150 mM NaCl. The shared $m_{D\text{-}N}$ value was fitted as 0.74 ± 0.07 kcal $mol^{-1}M^{-1}$.

Supplementary Tables

Supplementary Table S1. Stabilities of NCBD variants expressed as the urea concentration at 50% denaturation. The proteins were denatured with urea and the structural transition was monitored with CD at 222 nm at 277 K. The solutions were either 50 mM potassium formate buffer (pH 3.0), 150 mM NaCl or 20 mM sodium phosphate buffer (pH 7.4), 150 mM NaCl. The data were fitted to a two-state model in GraphPad Prism. The parameters for the signal of the native state, the denatured and native baseline slopes, and the $m_{D\text{-}N}$ value were shared among the data sets in the curve fitting to allow fitting of the low stability $NCBD_{D/P}^{ML}$ and $NCBD_{D/P}^{AltAll}$ variants. The shared $m_{D\text{-}N}$ value was fitted as 0.82 ± 0.025 kcal $mol^{-1}M^{-1}$. The errors are the standard error from the curve fitting and likely a low estimate of.

Variant	pH	$[Urea]_{50\%}$ (M)	$\Delta G_{D\text{-}N}$ (kcal/mol)
$NCBD_{D/P}^{AltAll}$	3.0	0.22 ± 0.04	0.18 ± 0.03
$NCBD_{D/P}^{ML}$	3.0	0.38 ± 0.04	0.31 ± 0.03
$NCBD_{D/P}^{ML}$	7.4	0.68 ± 0.04	0.55 ± 0.04
$NCBD_{D/P}^{T2073W}$	3.0	2.13 ± 0.04	1.73 ± 0.06
$NCBD_{D/P}^{T2073W}$	7.4	2.08 ± 0.04	1.70 ± 0.06
$NCBD_{Human}$	3.0	1.61 ± 0.04	1.31 ± 0.05
$NCBD_{Human}$	7.4	1.94 ± 0.05	1.58 ± 0.06
$NCBD_{Human}^{T2073W}$	3.0	3.24 ± 0.07	2.64 ± 0.10
$NCBD_{Human}^{T2073W}$	7.4	3.46 ± 0.08	2.82 ± 0.11

Supplementary Table 2. Rate constants and kinetic amplitudes of folding for different D/P NCBD Trp variants measured in temperature jump experiments. The experiments were conducted using 200 μM protein in 20 mM sodium phosphate (pH 7.4), 150 mM NaCl. The jump in temperature was from 277 to 285.5 K.

Variant	**Amplitude**	k_{obs} **(s^{-1})**
$NCBD_{D/P}^{T2073W}$	1.06 ± 0.06	2400 ± 200
$NCBD_{D/P}^{L2067W}$	0.91 ± 0.02	1530 ± 50
$NCBD_{D/P}^{S2078W}$	0.35 ± 0.01	700 ± 60
$NCBD_{D/P}^{H2107W}$	0.47 ± 0.02	820 ± 80
$NCBD_{D/P}^{Q2108W}$	0.73 ± 0.02	960 ± 60

Supplementary Table 3

	$NCBD_{D/P}^{ML}$	$NCBD_{D/P}^{AltAlt}$
NMR distance and dihedral restraints		
Distance restraints		
Total NOEs	329	347
Intra-residue	113	145
Sequential ($\|i-j\|=1$)	122	111
Medium-range ($1<\|i-j\|<4$)	76	76
Long-range ($\|i-j\|>5$)	18	39
Total dihedral angle restraints	76	63
$^{3}J_{HN\alpha}$ scalar couplings	39	43
$^{13}C^{\alpha}$ chemical shifts	100	100
Structure statistics		
Average CYANA target function value ($Å^2$)	2.5 ± 0.2	2.8 ± 0.3
Violations		
Distance constraints (>0.5 Å)	0	0
Dihedral angle constraints (>5°)	0	0
Deviations from idealized geometry		
Bond lengths (Å)	0	0
Bond angles (°)	0	0
Impropers (°)	0	0
Average pairwise r.m.s. deviation (Å)		
Backbone residues	2.2 ± 1.0	0.5 ± 0.2
Heavy atoms residues	2.7 ± 1.0	1.0 ± 0.3