Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/4/10/eaau4130/DC1

Supplementary Materials for

Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins

Per Jemth*, Elin Karlsson, Beat Vögeli, Brenda Guzovsky, Eva Andersson, Greta Hultqvist, Jakob Dogan, Peter Güntert, Roland Riek, Celestine N. Chi*

*Corresponding author. Email: chi.celestine@imbim.uu.se (C.N.C.); per.jemth@imbim.uu.se (P.J.)

Published 24 October 2018, *Sci. Adv.* **4**, eaau4130 (2018)
DOI: 10.1126/sciadv.aau4130

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/4/10/eaau4130/DC1)

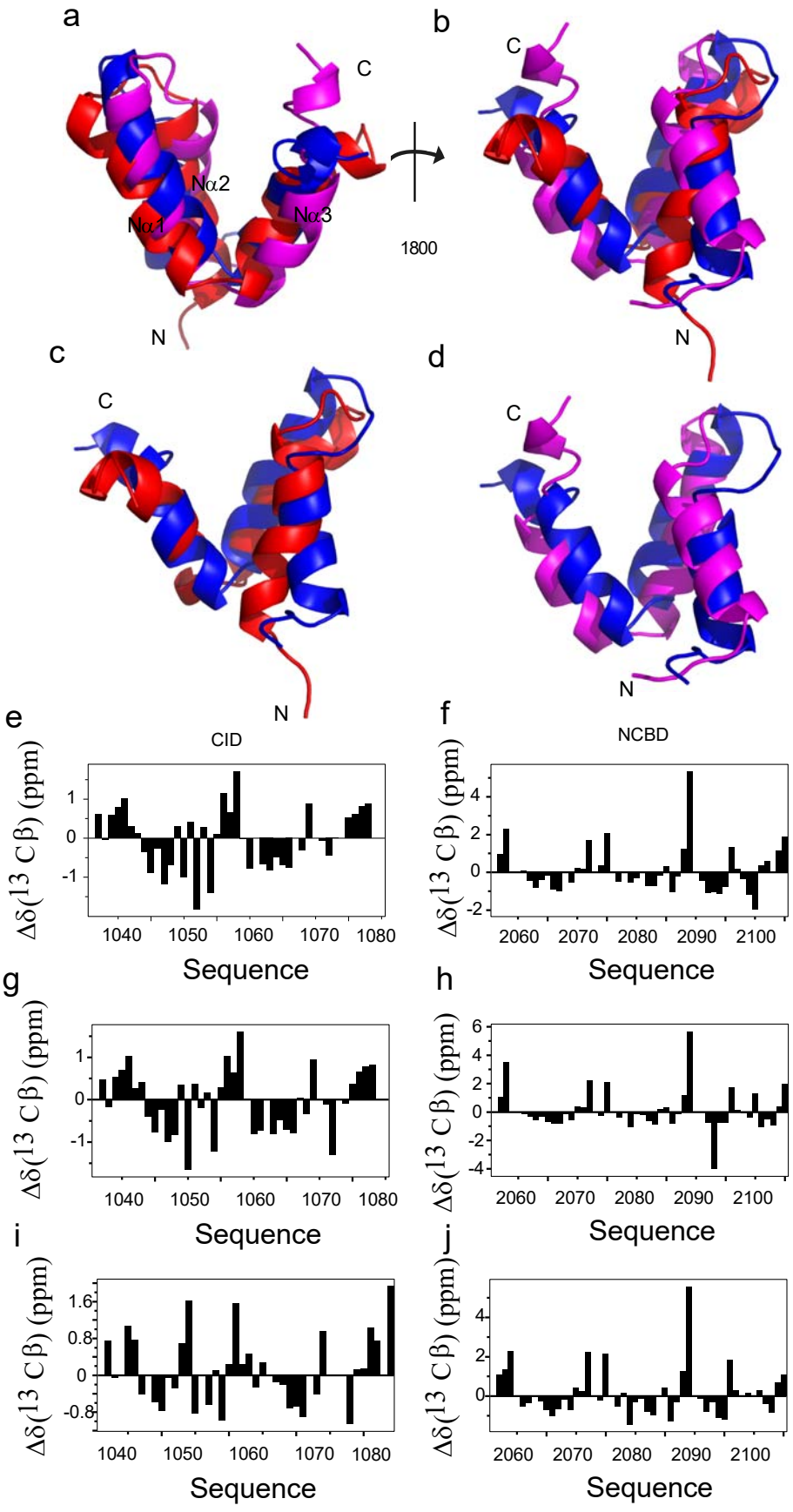

a
b
c
d
C
N
Nα1
Nα2
Nα3
1800
e
f
CID
NCBD
Δδ(13 Cβ) (ppm)
Sequence
g
h
i
j

Fig. S1. Structural changes taking place and chemical shift restraints. (a) Overlay of the most ancient Cambrian-like D/P NCBD (red), Ordovician-Silurian 1R/2R NCBD (blue) and extant human CREBBP NCBD (magenta) domains in complex with CID domains (1R CID for the older complexes and human NCOA3 CID for extant human, respectively). **(b)** The structures are reoriented by 180° to show the displacement of Nα3 in the most ancient complex. Also shown is the slight tilting of the Nα1 of human CREBBP NCBD as compared to the older NCBD domains. The orientations of the helices are also reflected in the RDCs (fig. S2). **(c)** Overlay of the most ancient Cambrian-like complex with the Ordovician-Silurian 1R/2R complex and **(d)** the 1R/2R complex with the extant human complex. **(e** to **j)** Plots of the difference of experimental Cβshifts to those of random coil values as a function of the amino acid sequences for ancestral and modern CID domains **(e)** and NCBD domains **(f).** Note the different scales on the y-axes in the different panels.

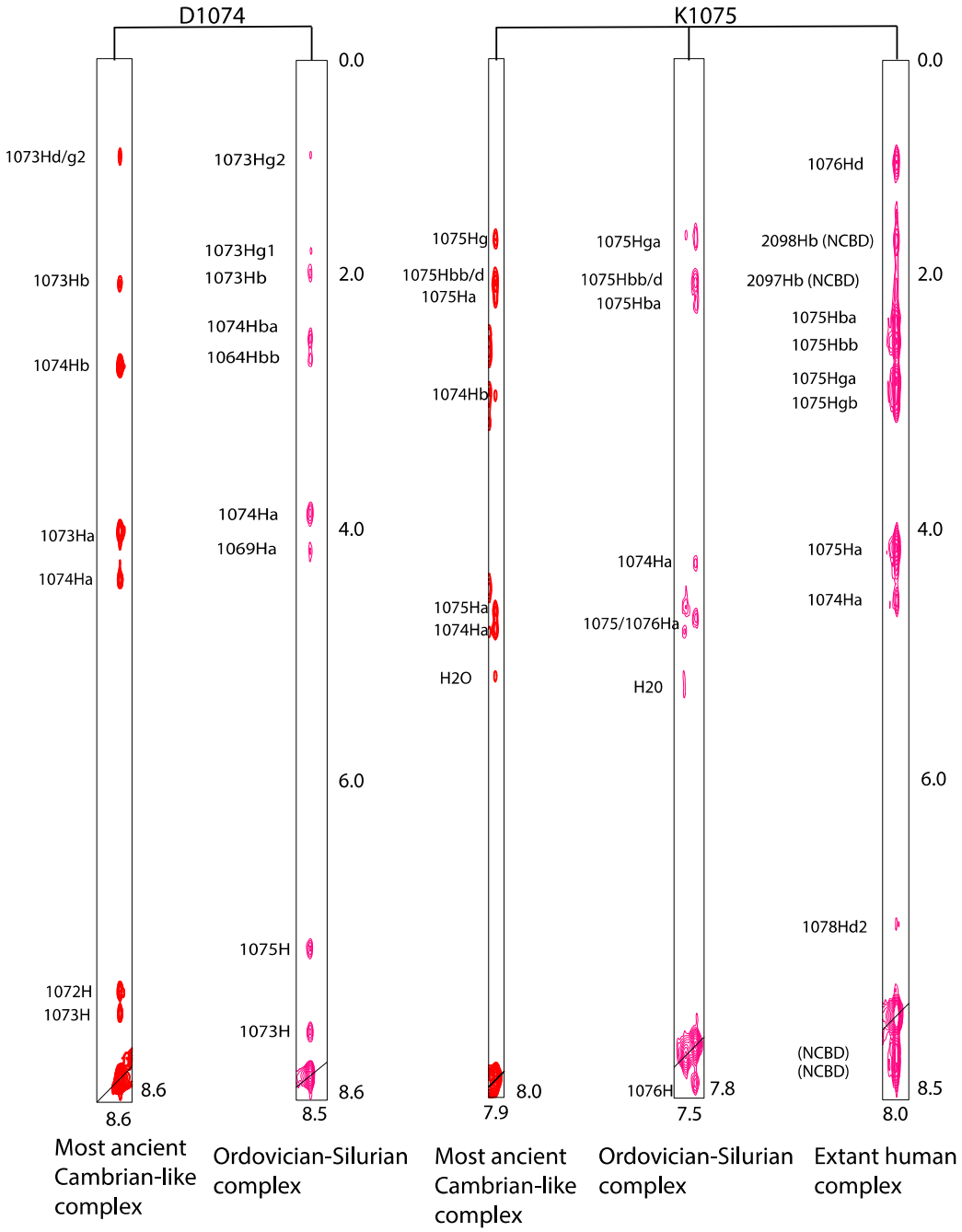


Fig. S2. Strips from [^{1}H-^{1}H]-NOESY-[^{1}H-^{15}N]-HSQC spectra showing resonances emanating from residues 1074 and 1075 of the respective CID domains. Cross peaks arising from inter- and intra-molecular contacts are shown. New contacts appear in the evolutionarily younger complexes. Note that human NCOA3 has a proline at position 1074 and thus contains no H-N.

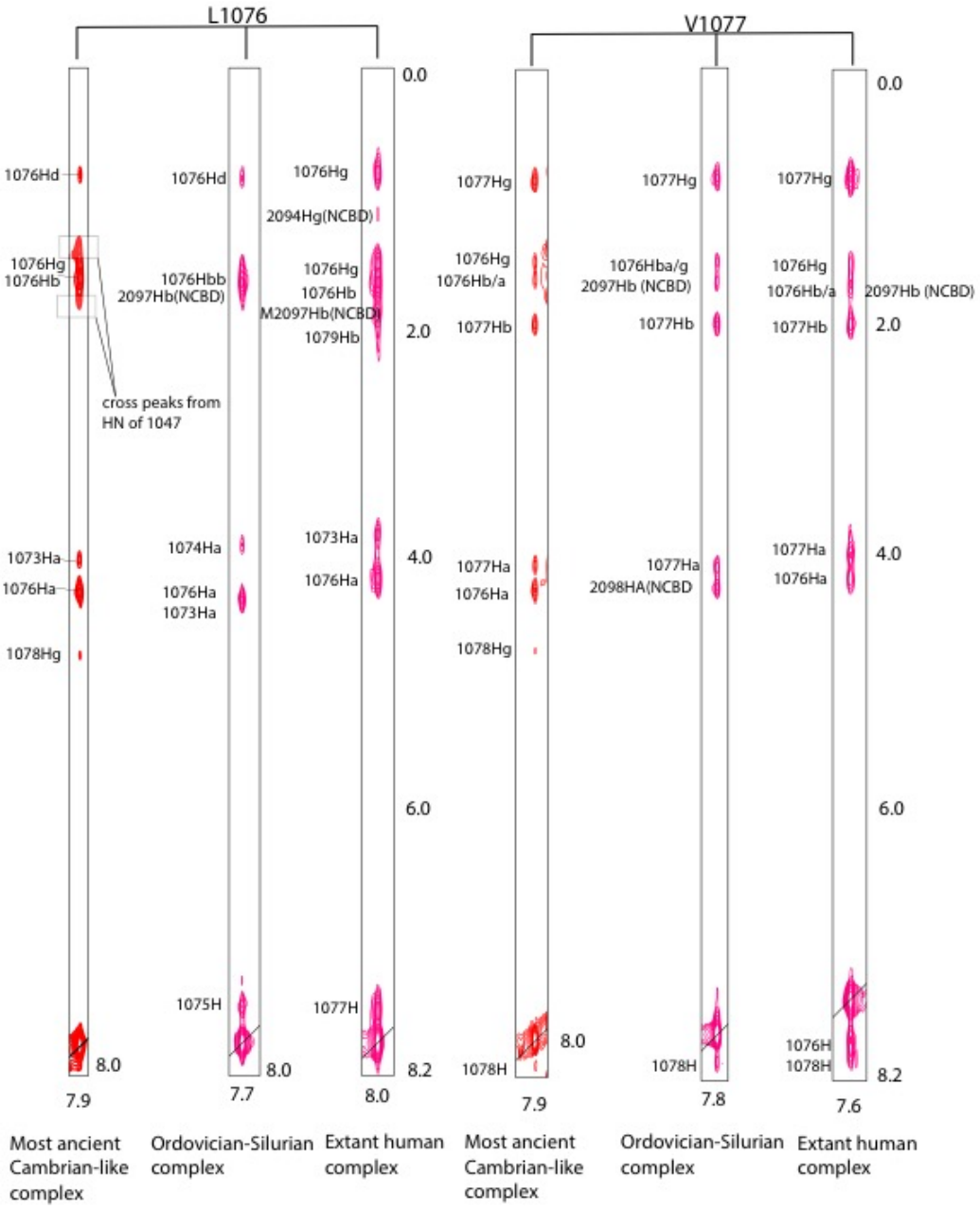


Fig. S3. Strips from [^{1}H-^{1}H]-NOESY [^{1}H-^{15}N]-HSQC spectra showing resonances emanating from residues 1076 and 1077 of the CID domains. Cross peaks arising from inter- and intra-molecular contacts are shown. New contacts appear in the evolutionarily younger complexes.

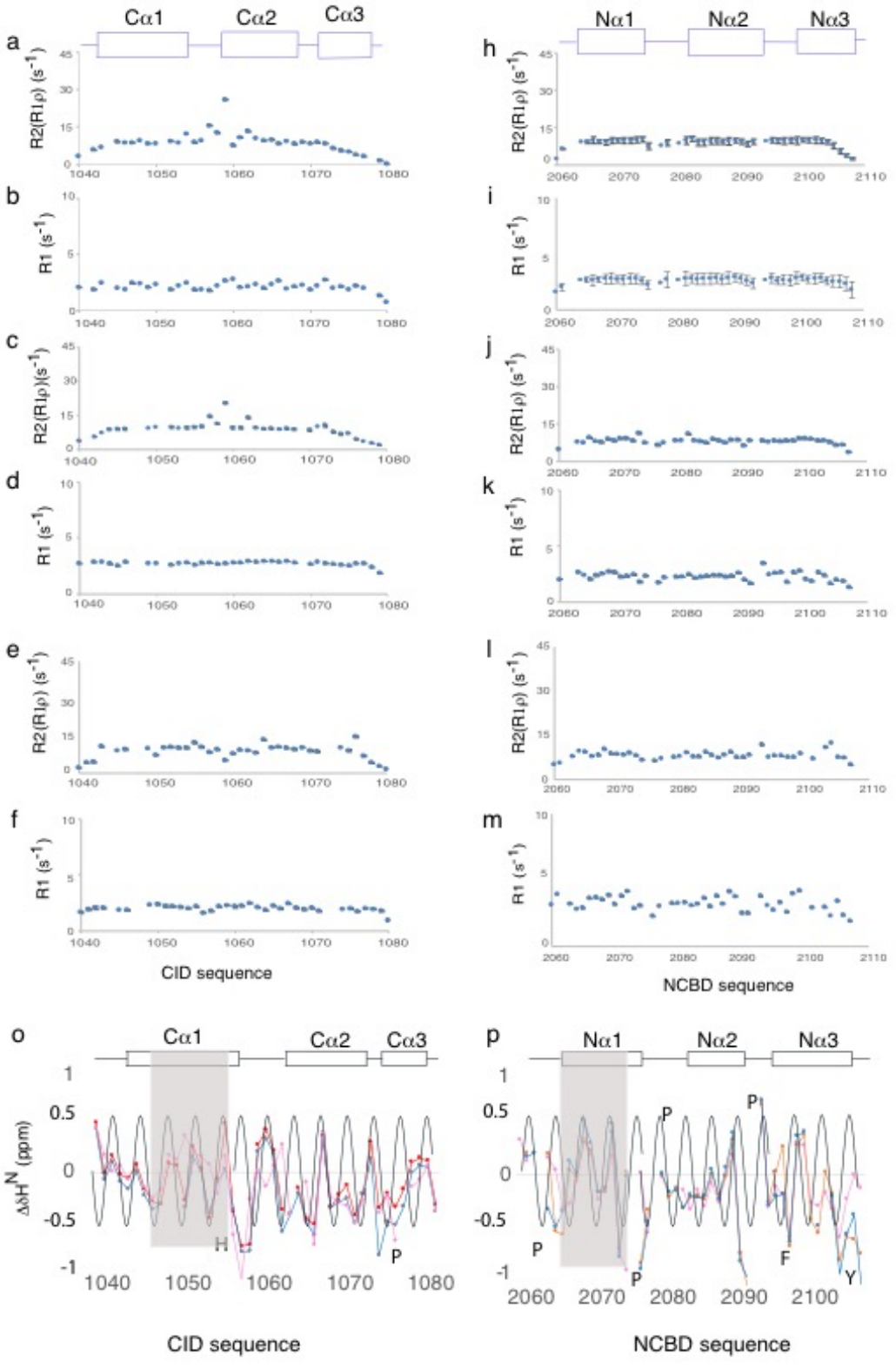
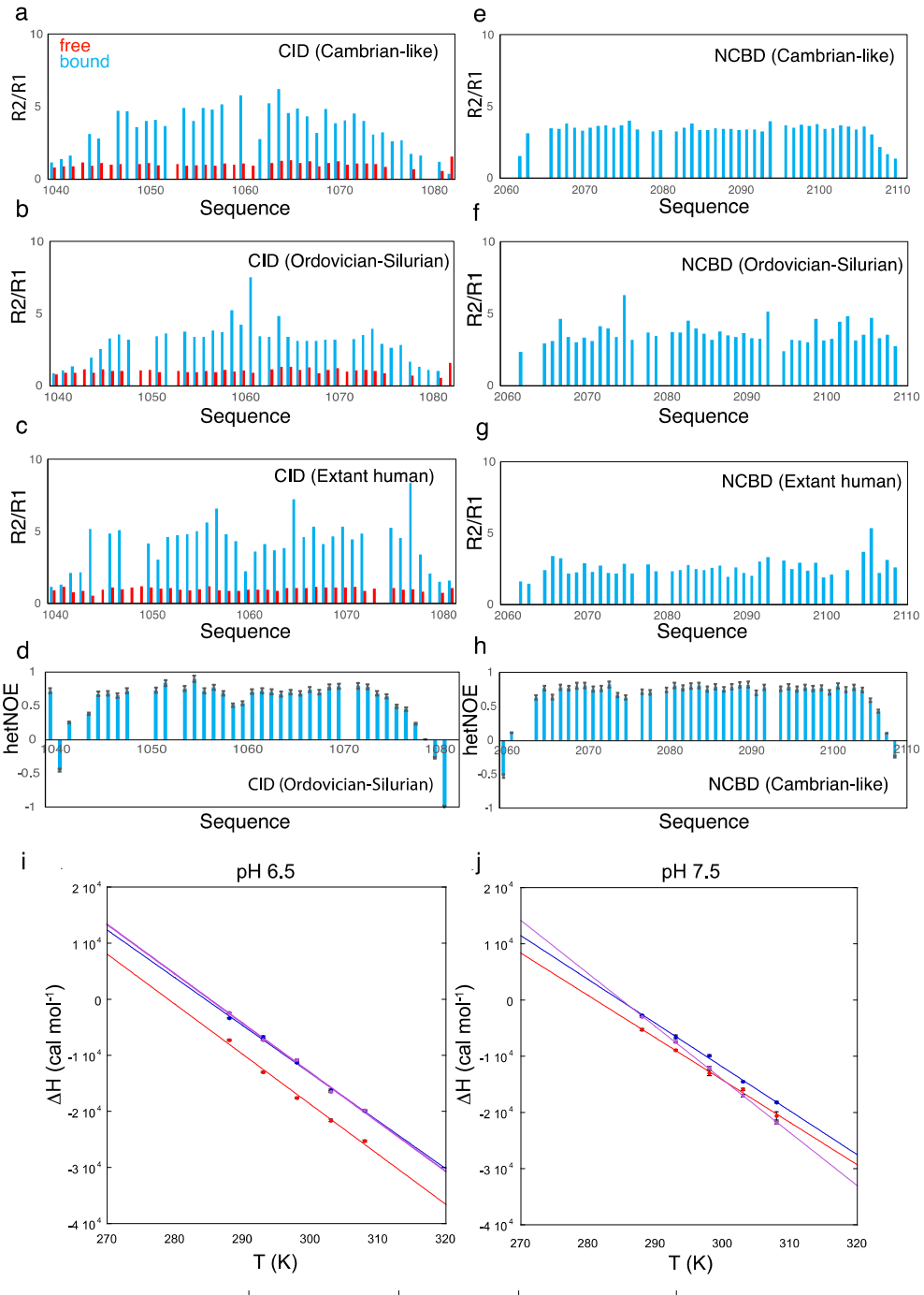

Cα1
Cα2
Cα3
Nα1
Nα2
Nα3
a
b
c
d
e
f
h
i
j
k
l
m
R2(R1ρ) (s-1)
R1 (s-1)
CID sequence
NCBD sequence
o
p
ΔδHN (ppm)
H
P
F
Y

Fig. S4. ^{15}N relaxation data for the bound CID and NCBD domains and amide secondary chemical shifts of the complexes. Relaxation rates plotted versus amino acid sequence for the CID domains **(a** to **f)** and NCBD domains **(h** to **m)** for the three complexes. (a to c) the most ancient Cambrian-like complex. **(a)** Transverse ^{15}N R_2 derived from $R_{1\rho}$ and R_1 relaxation rates plotted as a function of the amino acid sequence. The spectra were recorded with a spin lock frequency of 2 kHz, thus the contributions of all exchange processes slower than 80 μs (present in $R_{2\beta}$) were suppressed. **(b)** Longitudinal ^{15}N R_1 relaxation rate (reports on ps-ns motions) plotted as a function of the amino acid sequence. **(c** and **d)** The Ordovician-Silurain 1R/2R complex; see legend (a and b). **(e** and **f)** The extant human complex; see legend for (a-b). The experiments depicted in (a to f) were performed with labeled CID domains. In panels **(h** to **n)** similar experiments as in panels (a to f) were performed with labeled NCBD domains. Errors are propagated errors from single exponential fits. **(o)** and **(p)** Amide secondary chemical shifts of the complexes are indicative of bent helices**.** The differences of experimental amide shift from the random coil value ($\Delta\delta$ H^N) are plotted against the amino acid sequence for the respective CID/NCBD complexes. Red, blue and magenta connected dots are for the most ancient Cambrian-like, Ordovician-Silurian 1R/2R and extant human complex, respectively. The black lines symbolize the wave-like pattern of the helices. The positive shifts are a result of short hydrogen bonds at the concave side while the negative shifts are from long hydrogen bonds on the convex side. There are 3-4 residues per repeat giving the wave-like signature which is characteristic of bent helices. The grey shading shows regions of increased bending seen in the structure, indicating that the bending is not an artefact from calculation. Also indicated are the positions of Pro and aromatic residues, since the ring current from these residues can change amid proton shifts



pH	Thermodynamic parameter	Cambrian-like	Ordovician-Silurian	Extant Human
7.5	ΔC_p (cal mol^{-1} K^{-1})	-740 ± 30	-780 ± 20	-940 ± 10
6.5	ΔC_p (cal mol^{-1} K^{-1})	-890 ± 60	-850 ± 40	-880 ± 50

Fig. S5. Backbone subnanosecond motions and ΔC_p for the three CID/NCBD complexes. Ratio of transverse and longitudinal relaxation rates plotted versus amino acid sequence for the CID domains **(a** to **c)** and NCBD domains **(f** and **g)** for the three complexes. Free (red) and bound (light-blue) for the **(a)** Cambrian-like complex, **(b)** the Ordovician-Silurian 1R/2R complex, and **(c)** the extant human complex. Data for CID-bound NCBD domains are shown in **(e, f and g**). **(d** to **h)** [^{1}H]-^{15}N heteronuclear NOE rates. Negative and zero values indicate large H-N bond amplitudes of motions faster than 1 ns. Errors are propagated errors from single exponential fits. ΔH was measured at different temperatures and the slope corresponds to ΔC_p. **(i)** pH 6.5 and **(j)** pH 7.5.

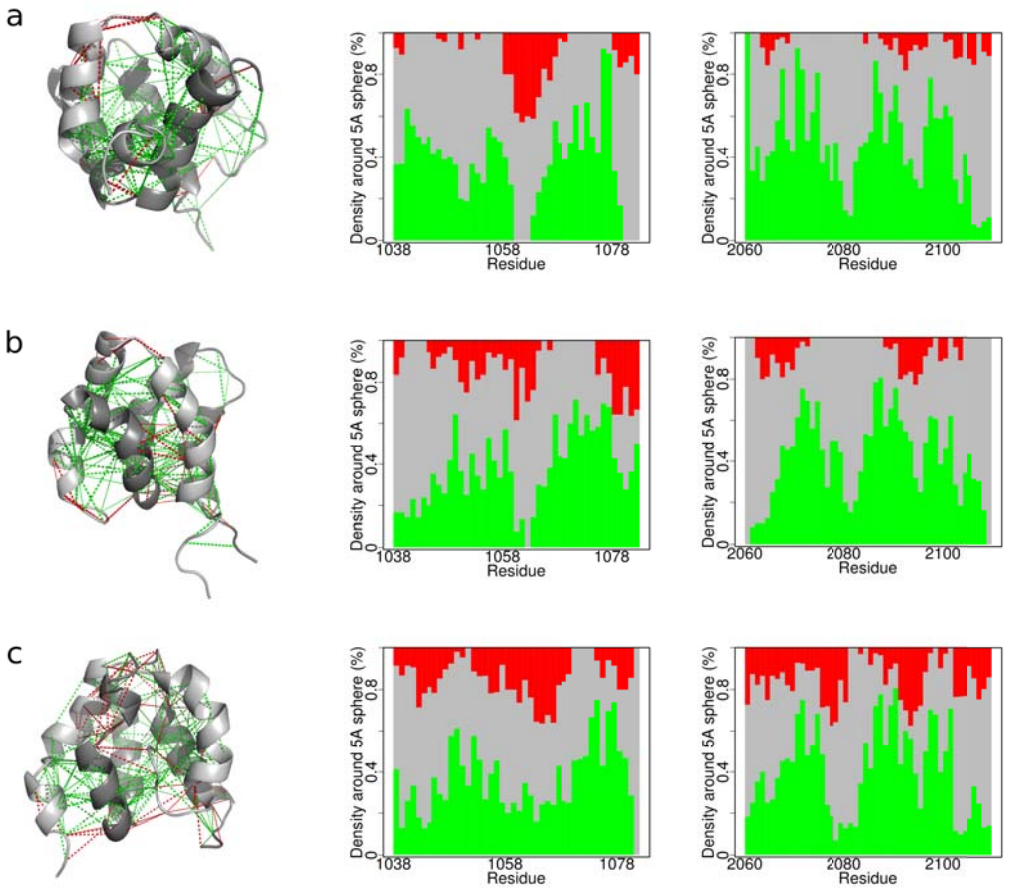


Fig. S6. Frustration in the CID/NCBD complexes without electrostatics. Mutational frustration for each complex, calculated using electrostatics=0. Left panel: CID/NCBD backbones are displayed as grey ribbons (darkest grey is NCBD), the direct contacts with solid lines and the water-mediated interactions with dashed lines. Minimally frustrated interactions are shown in green, highly frustrated contacts in red, and neutral contacts are not drawn. Middle and Right panel: Histograms of local frustration plotted as a function of amino acid sequence. The proportion of contacts within 5 Å of the Cα atom of each residue is plotted, classified according to their frustration index. **(a)** The Cambrian-like complex contains the 1R CID (middle panel) and D/P NCBD (right panel) domain. **(b)** The Ordovician-Silurian complex contains 1R CID (middle panel) and 1R/2R NCBD (right panel) domain. **(c)** The Extant human complex contains NCOA3 CID (middle panel) and CREBBP NCBD (right panel) domain.

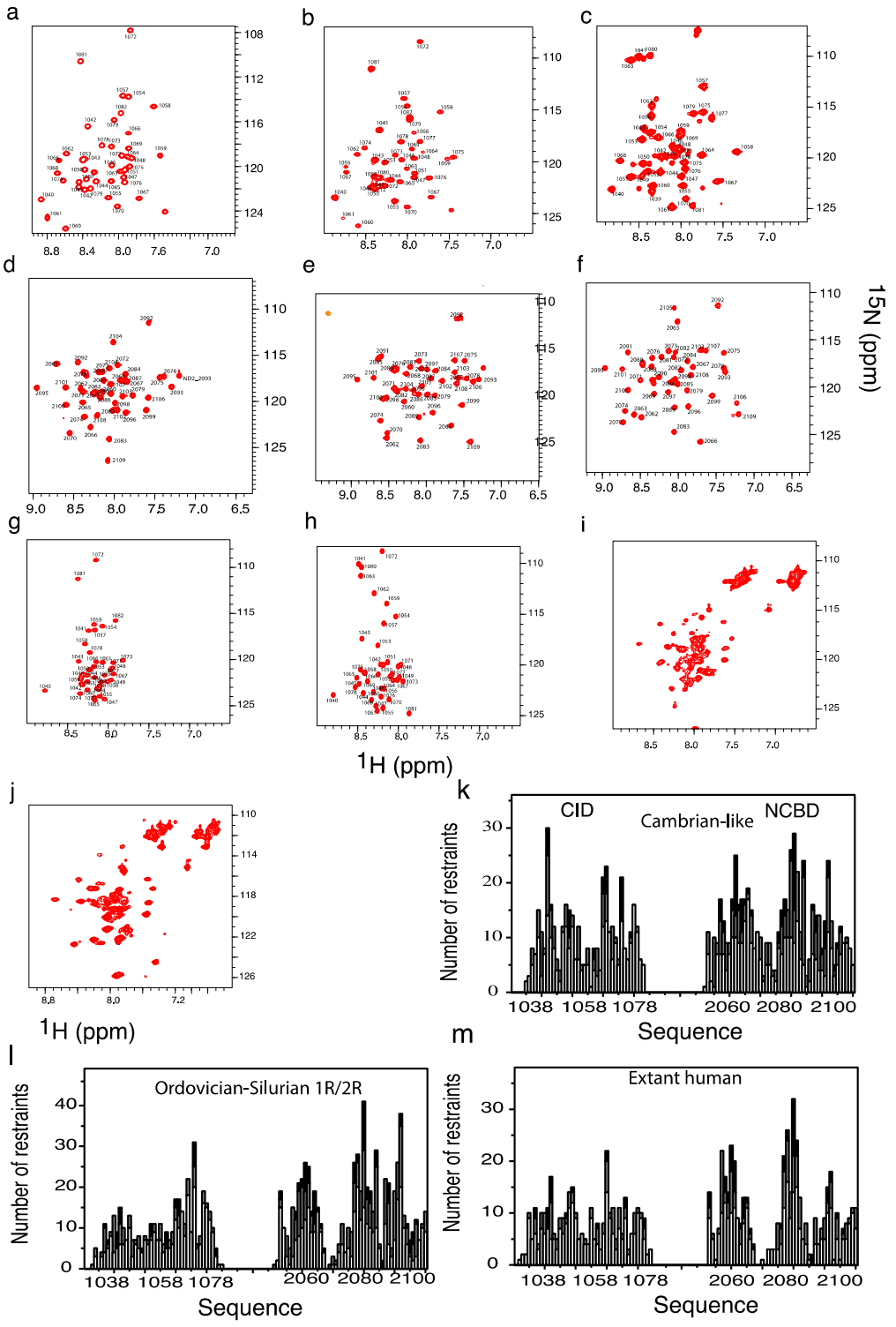

a
b
c
d
e
f
g
h
i
j
k
l
m
15N (ppm)
1H (ppm)
1H (ppm)
CID
Cambrian-like
NCBD
Number of restraints
Sequence
Ordovician-Silurian 1R/2R
Extant human

Fig. S7. [^{1}H-^{15}N]-HSQC correlation spectra for bound and free CID and NCBD domains and the number of restraints in the structure determination as a function of amino acid residue. (a) Most ancient Cambrian-like complex: ^{15}N-labeled 1R CID bound to unlabeled D/P NCBD. **(b)** Ordovician-Silurian complex: ^{15}N-labeled 1R CID bound to unlabeled 1R/2R NCBD. **(c)** extant human complex: ^{15}N-labeled NCOA3 CID bound to unlabeled human CREBBP NCBD. **(d)** Most ancient Cambrian-like complex: ^{15}N-labeled D/P NCBD bound to unlabeled 1R CID. **(e)** Ordovician-Silurian complex: ^{15}N-labeled 1R/2R NCBD bound to unlabeled 1R CID. **(f)** extant human complex: ^{15}N-labeled human CREBBP NCBD bound to unlabeled human NCOA3 CID. **(g)** ^{15}N-labeled 1R CID. **(h)** ^{15}N-labeled NCOA3 CID. **(i)** ^{15}N-labeled D/P NCBD. **(j)** ^{15}N-labeled CREBBP NCBD. **(k)** The most ancient Cambrian-like complex, **(l)** the Ordovician-Silurian 1R/2R complex and **(m)** the extant human complex. Sequential (white), medium (grey) and long range (black) distances.

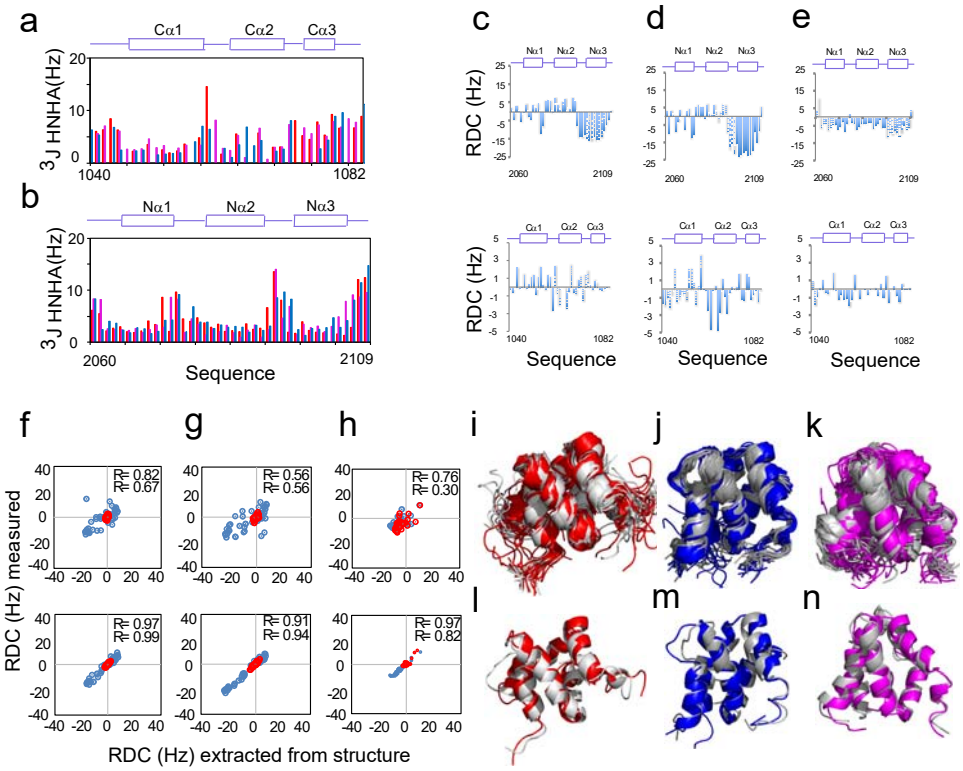

a
Cα1
Cα2
Cα3
3J HNHA(Hz)
1040
1082
b
Nα1
Nα2
Nα3
2060
2109
Sequence
c
d
e
RDC (Hz)
f
g
h
i
j
k
l
m
n
R= 0.82
R= 0.67
R= 0.56
R= 0.56
R= 0.76
R= 0.30
R= 0.97
R= 0.99
R= 0.91
R= 0.94
R= 0.97
R= 0.82
RDC (Hz) measured
RDC (Hz) extracted from structure

Fig. S8. Restraints used for structure calculations and refinement. Three bond HNHA couplings were determined as described in the methods section, for **(a)** the CID domain and **(b)** the NCBD domain for most ancient Cambrian-like (red), Ordovician-Silurian 1R/2R (blue) and extant human (magenta) CID/NCBD complexes. The coupling constants were then converted into φ dihedral angles following the relationship described by Vuister and Bax[1]. All three helices for the NCBD and the first two helices of the CID domains showed the characteristic small couplings (1-4 Hz) associated with α helices. The loop regions as well as the Cα3 showed higher values (5-10 Hz). For the RDC measurements**,** the CID/NCBD complexes were aligned in pf1 phage and ^{1}D H-N RDCs were determined in an IPAP manner as described in the method section. RDCs for the most ancient Cambrian-like complex **(c)**, Ordovician-Silurian 1R/2R complex **(d)** and human extant complex **(e)**. In the top row, RDCs as function of residue number for the NCBD domains are shown, while in the bottom row are those for the CID domains. For structural refinement, RDCs were extracted from the structures that were calculated without RDCs (x-axis) and plotted against the measured RDCs (y-axis) (top panels). The RDCs were then included in the structure calculations and the RDCs from these structures were later extracted and plotted against the measured RDCs (bottom panel). The R-value (indicated) shows how well the extracted values agree with the measured values. **(f)** Most ancient Cambrian-like complex, **(g)** Ordovician-Silurian 1R/2R complex and **(h)** extant human complex. Blue are RDCs for the NCBD domain and red for the CID domain (note that the alignment in the pf1 phages were much stronger for the NCBD than for the CID domains). **(i** to **n)** Representation of the complexes calculated without (light grey) and with RDCs (colored) for **(i** and **l)** most ancient Cambrian-like complex **(j** and **m)** Ordovician-Silurian 1R/2R complex and **(k** and **f)** extant human complex. For clarity, only two of the structures are shown in **(l-f).**

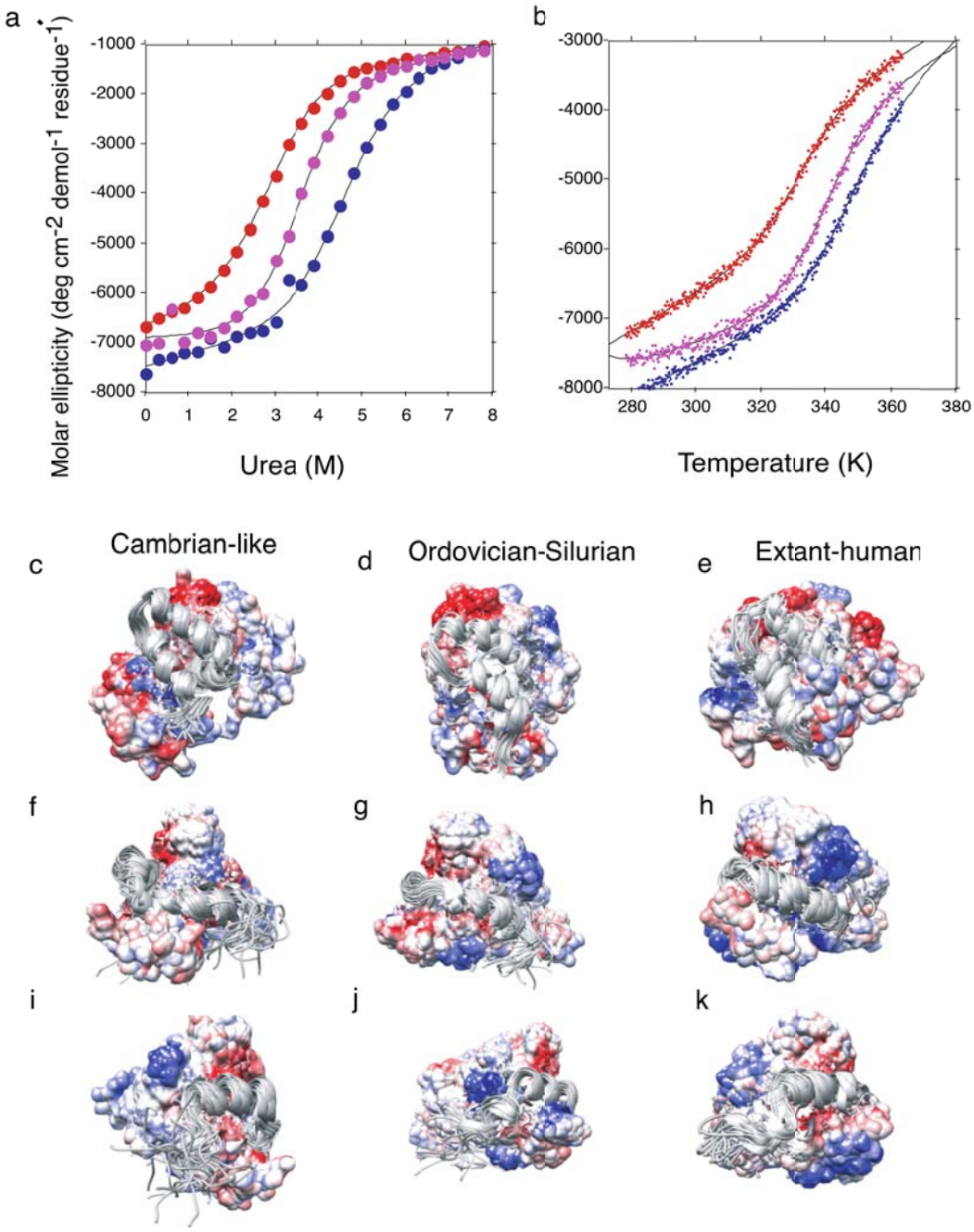

a
b
Molar ellipticity (deg cm-2 demol-1 residue-1)
-1000
-2000
-3000
-4000
-5000
-6000
-7000
-8000
0
1
2
3
4
5
6
7
8
Urea (M)
-3000
-4000
-5000
-6000
-7000
-8000
280
300
320
340
360
380
Temperature (K)
c
Cambrian-like
d
Ordovician-Silurian
e
Extant-human
f
g
h
i
j
k

Fig. S9. Urea and temperature denaturation experiments and Coulomb surfaces of the ancestral and extant CID/NCBD complexes. The Molar ellipticity is plotted as function of **(a)** urea concentration in units of molar (M) and **(b**) temperature in Kelvin (K) for the Cambrian-like (red), Ordovician-Silurian (blue) and extant human (magenta) CID/NCBD complexes, at 40 μM CID and 15 μM NCBD, respectively. The data were fitted to the standard chemical and thermal denaturation equations, respectively, for global two-state unfolding. The cooperativity of the urea denaturation curves denoted the $m_{D\text{-}N\ value}$ is similar for all three complexes within error (0.95 ± 0.04, 0.92 ± 0.12 and 1.06 ± 0.12 kcal $mol^{-1}M^{-1}$, respectively). The urea midpoint was lower for the Cambrian-like (2.9 ± 0.04 M), compared to the Ordovician-Silurian (4.6 ± 0.12 M) and the human (3.5 ± 0.09 M) complexes. These results were also mirrored in the thermal denaturation experiments with T_m values of 325 K, 344 K and 340 K, respectively. Coulomb surfaces color code: red < -10 kcal/(mol)(e), blue >+10 kcal/(mol)(e). In (c), (d) and (e) the NCBD domains are displayed in grey cartoons while the CID domains are shown as surfaces. In (f), (g) and (h) the CID domains are displayed in grey cartoons and the NCBD domains as surfaces. Panels (i), (j) and (k) are as (f), (g) and (h) but the complexes are reoriented to show the Cα3 helix of the CID domains.

Table S1. ITC parameters and intermolecular NOEs. Tab 1. pH dependence, Tab 2 temperature variation at pH 7.5, Tab 3 temperature variation at pH 6.5 and Tab 4 intermolecular NOEs. Extra intermolecular NOEs, which are not present in the Cambrian-like complex are highlighted in yellow.