ChemMedChem

Supporting Information

A Step toward NRF2-DNA Interaction Inhibitors by Fragment-Based NMR Methods

Sven Brüschweiler,* Julian E. Fuchs, Gerd Bader, Darryl B. McConnell, Robert Konrat, and Moriz Mayer

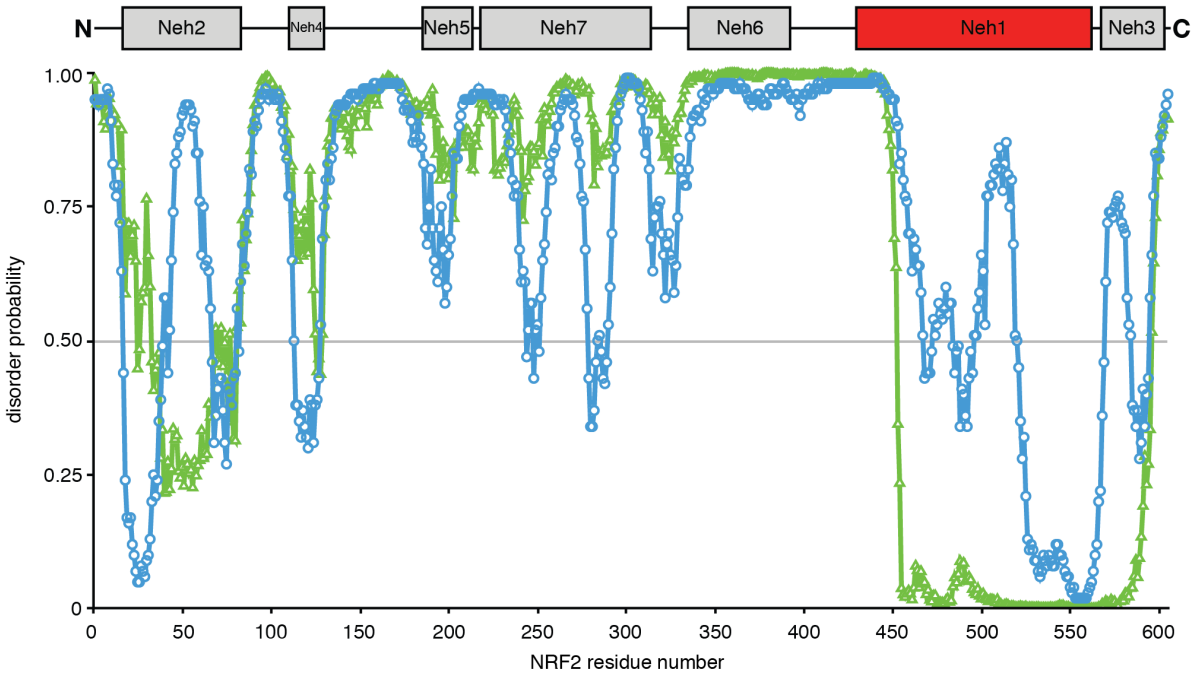
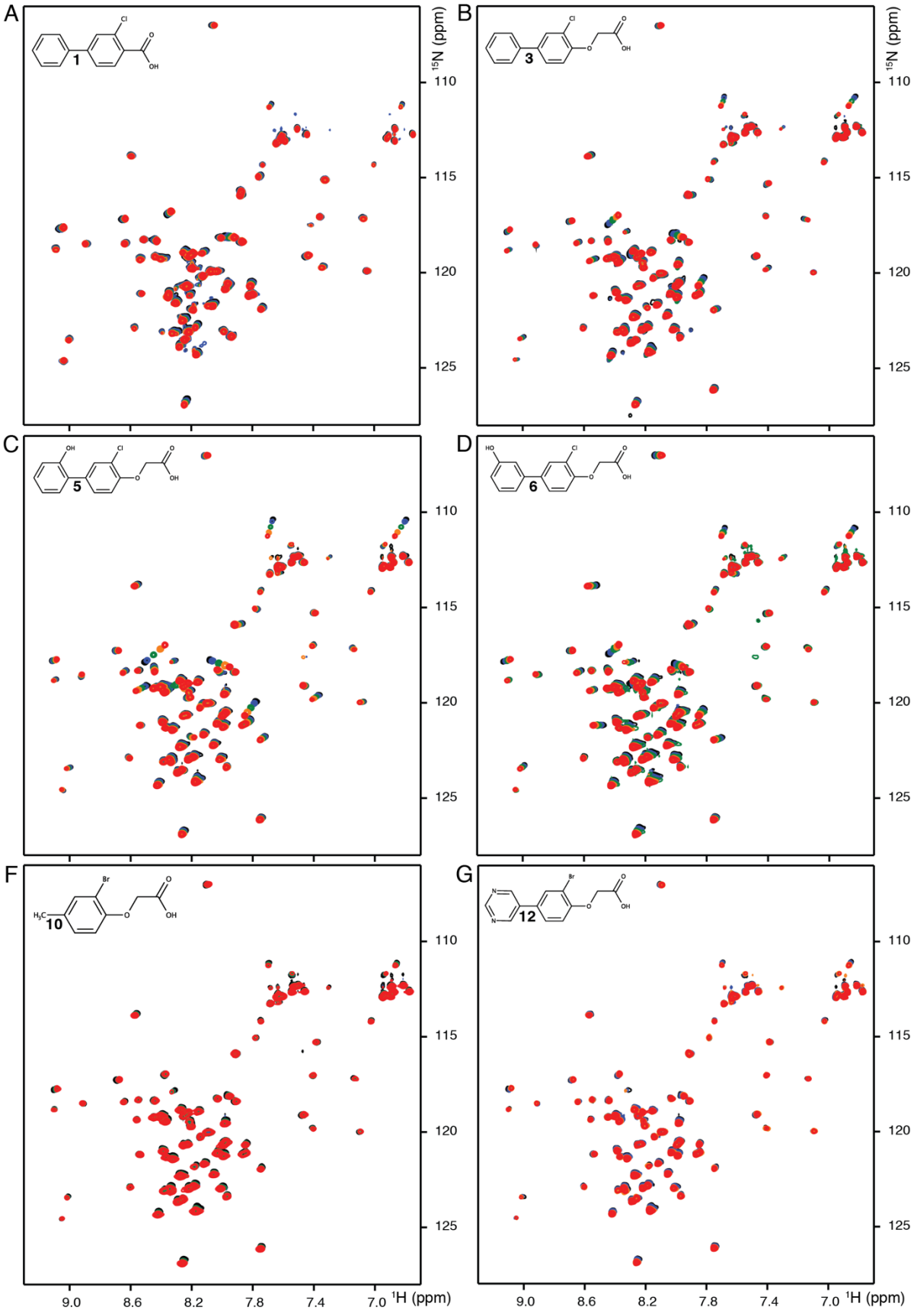


Figure S1. Predicted disorder probabilities for human NRF2. Disorder prediction using DISOPRED3[1] (blue) and NetSurfP-2.0[2] (green) with the schematic NRF2 domain structure on top. Compared to the Neh1 domain structure (PDB ID 7O7B) determined in this article the DISOPRED3 webserver overestimates the disorder of the CNC domain, whereas the NetSurfP-2.0 webserver correctly predicts the CNC domain, however, it wrongly predicts the basic region to be folded over its full length in solution.


A
B
C
D
F
G
1
3
5
6
10
12
^{15}N (ppm)
110
115
120
125
9.0
8.6
8.2
7.8
7.4
7.0
^{1}H (ppm)

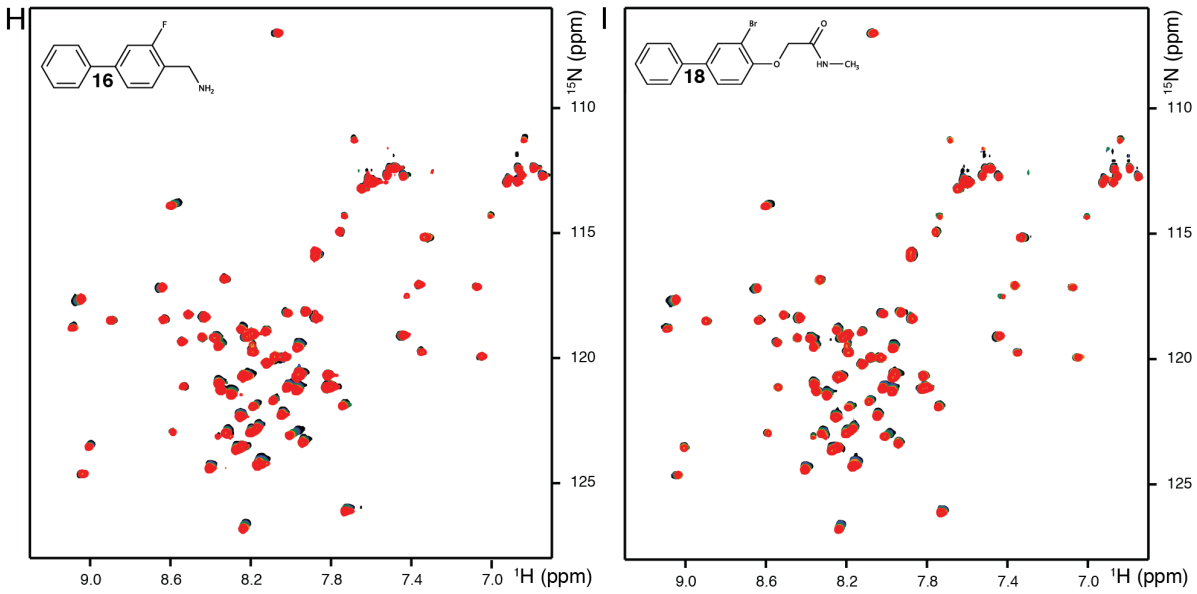


Figure S2. NMR-based fragment screen identified Neh1 binders. (A-I) Overlay of ^{1}H-^{15}N HSQC spectra of uniformly labeled 50 μM ^{15}N-Neh1-ΔLZIP (red) titrated with 250, 750, 1500, and 2000 μM of respective compound (from orange to black) (A) compound **1**. (B) compound **3**. (C) compound **5**. (D) compound **6**. (E) compound **10**. (F) compound **12**. (H) compound **16**. (I) compound **18**.

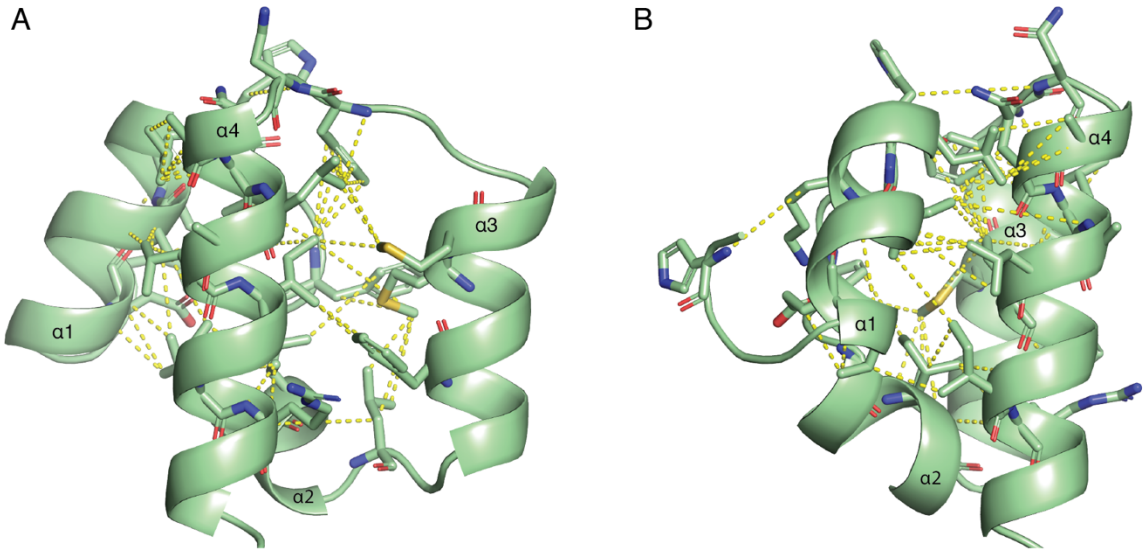


Figure S3. Long range NOE distance restraints. (A-B) The unambiguous long range distance restraints (i-j≥5) are plotted as yellow lines onto the Neh1-ΔLZIP lowest energy structure (PDB ID 7O7B). Residues H453-K506 of Neh1-ΔLZIP are shown and the rest of the protein is omitted for clarity. Residues for which long range NOEs were measured are shown in stick representation with the distance restraints drawn between heavy atoms to reduce the number of lines.

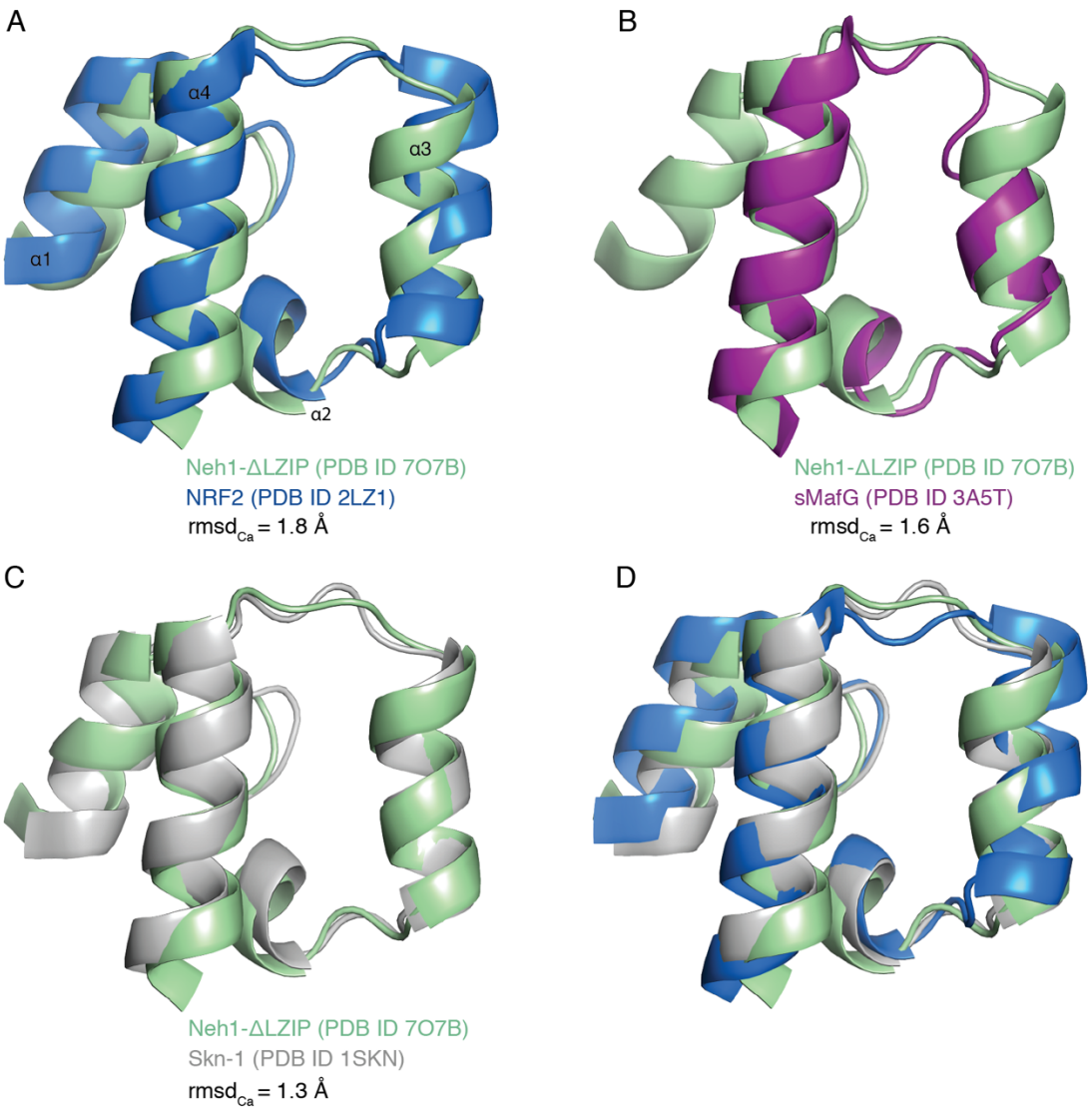


Figure S4. Overlay of the helical region of the NMR structure of Neh1-ΔLZIP (PDB ID 7O7B) and previously solved transcription factors, NRF2 (residues 445-523, PDB ID 2LZ1), sMafG (PDB ID 3A5T), and Skn-1 (PDB ID 1SKN). (A) Superimposition of the NMR structure of Neh1-ΔLZIP (PDB ID 7O7B) with the NMR structure of NRF2 (PDB ID 2LZ1), residues R456-G505 are shown for both structures and the disordered residues are omitted for clarity. (B) Superimposition of the NMR structure of Neh1-ΔLZIP (residues 456-505 are shown, PDB ID 7O7B) with the sMafG:ARE complex X-ray structure (residues 26-60 are shown and the rest of the protein is omitted for clarity, PDB ID 3A5T). (C) Superimposition of the NMR structure of Neh1-ΔLZIP (residues 456- 505 are shown, PDB ID 7O7B) with the Skn-1:DNA complex X-ray structure (residues 460-509 are shown and the rest of the protein is omitted for clarity, PDB ID 1SKN). (D) Superimposition of Neh1-ΔLZIP (PDB ID 7O7B) with NRF2 (PBD ID 2LZ1), and Skn-1 (PDB ID 1SKN).

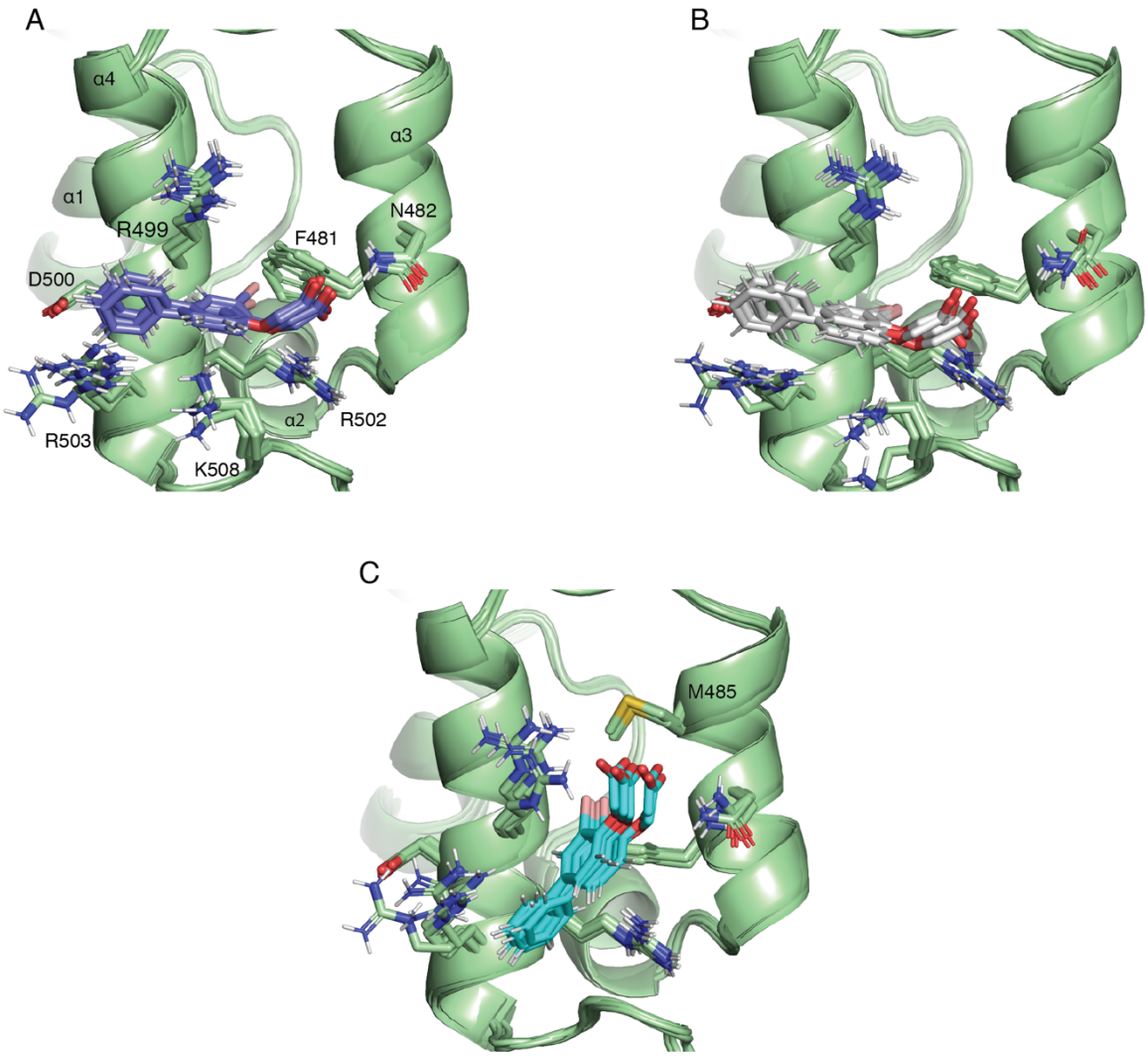


Figure S5. Data-driven docking structures of fragment **2**; ribbon representation of NRF2 (residues R456-A510 are shown and the rest of the protein is omitted for clarity, PDB ID 7O7B) complexed to compound **2**. (A) Shown in sticks representation are 5 representative conformations of **2** in cluster 1. (B) Shown in sticks representation are 5 representative conformations of **2** in cluster 2. (C) Shown in sticks representation are 5 representative conformations of **2** in cluster 3.

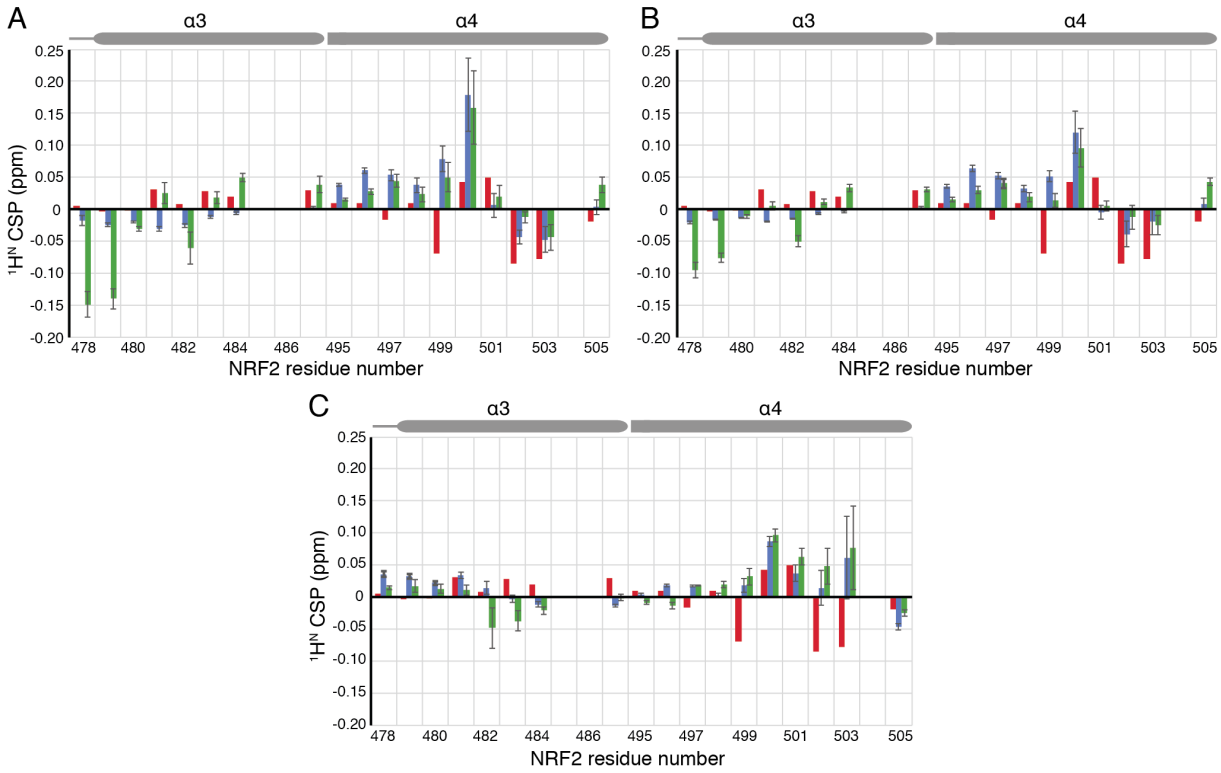


Figure S6. Experimental and simulated H^N chemical shift perturbations induced by fragment **2**. (A) Shown in red are the measured CSP values; the simulated values for the ring current (RC) effect (blue), and the sum of the RC and electric field (EF) effect (green) are shown as the average over 5 structures in cluster 1, which are plotted in Figure S5A. With the error being the standard deviation (SD) of the mean of the simulated CSP values. (B) Shown in red are the measured CSP values; the simulated values for the RC effect (blue), and the sum of the RC and EF effect (green) are shown as the average over 5 structures in cluster 2 shown in Figure S5B. With the error being the SD of the mean of the simulated CSP values. (C) Shown in red are the measured CSP values; the simulated values for the RC effect (blue), and the sum of the RC and EF effect (green) are shown as the average over 5 structures in cluster 3 shown in Figure S5C. With the error being the SD of the mean of the simulated CSP values.

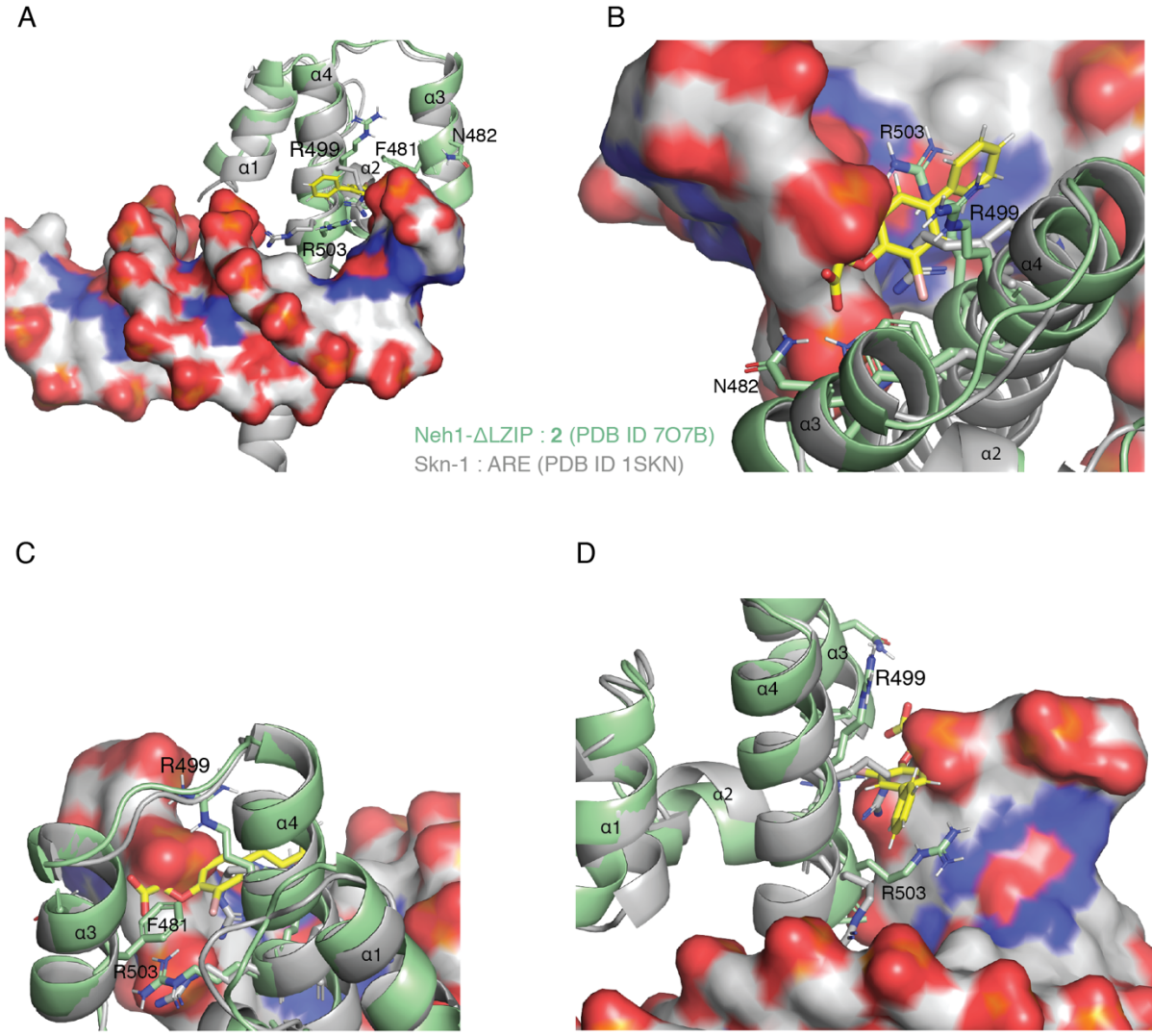


Figure S7. Superimposition of the helical region of the NMR Neh1-ΔLZIP:**2** docking structure and the X-ray structure of transcription factor Skn-1 in complex with DNA (PDB ID 1SKN). The superimposition further corroborates the overlapping NRF2 binding sites for small-molecule binders, described in this article, and ARE DNA. (A-D) Superimposition of NMR docking structure of Neh1-ΔLZIP:**2** (PDB ID 7O7B) with Skn-1:DNA complex X-ray structure (PDB ID 1SKN). Shown in yellow sticks representation is compound **2** in cluster 2 and the DNA of the Skn1:DNA complex is shown in molecular surface presentation. (A) In cartoon representation residues L454-K506 of Neh1-ΔLZIP, and residues G456-R525 of Skn-1:DNA and the rest of the proteins is omitted for clarity. Superimposition shows the partial overlap of binding sites between **2** and DNA and further underlines the importance of R499 and R503 in both compound **2** and DNA binding. (B-C) Detailed view of molecular interactions that show similar positioning of one of the DNA backbone phosphate groups and the phenoxy-acetic acid motif of **2**. (D) Detailed view of the compound binding pose suggests that higher affinity binders could potentially among others prevent the rearrangement of R499 that occurs upon DNA binding.

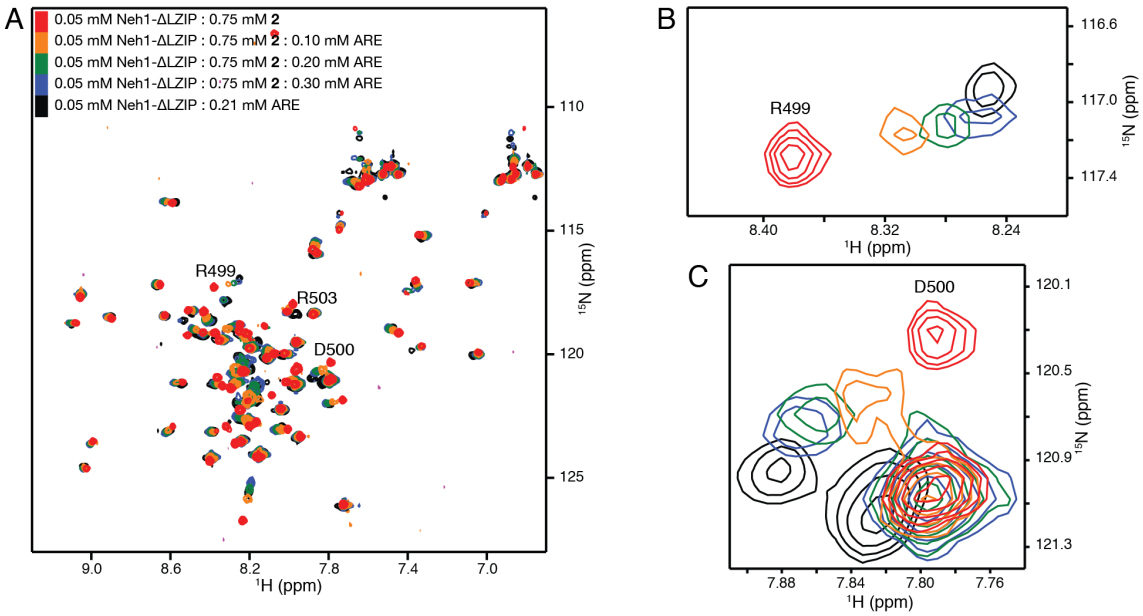


Figure S8. NMR competition experiment shows specific replacement of compound **2** with ARE DNA. (A-C) Overlay of ^{1}H-^{15}N HSQC spectra of uniformly labeled 50 μM ^{15}N-Neh1-ΔLZIP in the presence of 750 μM compound **2** (red) titrated with 100, 200, and 300 μM ARE DNA (from orange to blue), as reference for a DNA bound Neh1 state the ^{1}H-^{15}N HSQC spectrum of uniformly labeled 50 μM ^{15}N-Neh1-ΔLZIP in the presence of 210 μM ARE DNA is shown in black. (B-C) Selected regions of the ^{1}H-^{15}N HSQC spectra overlay showing close up views of R499 and D500 and their concentration-dependent chemical shift changes upon increasing amounts of ARE.

References

1.Jones, D. T.; Cozzetto, D., DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics* **2015,** *31* (6), 857-863.

2.Klausen, M. S.; Jespersen, M. C.; Nielsen, H.; Jensen, K. K.; Jurtz, V. I.; Sonderby, C. K.; Sommer, M. O. A.; Winther, O.; Nielsen, M.; Petersen, B.; Marcatili, P., NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. *Proteins-Structure Function and Bioinformatics* **2019,** *87* (6), 520-527.