Nucleic Acids Research, 2025, **53**, gkaf439
https://doi.org/10.1093/nar/gkaf439
Molecular and Structural Biology

The role of multivalency in the association of the eight twenty-one protein 2 (ETO2) with the nucleosome remodeling and deacetylase (NuRD) complex

Glory Dan-Dukor[1], Shengzhe Shang[2], Gage O. Leighton[3], Christopher R. Travis[4], Timothy D. Schwochert[3,4], Parnika Agrawal[3], Oyindamola Ajasa[3], Torry Li[3], Marcey L. Waters[4], Gordon D. Ginder[2,5], David C. Williams, Jr.[3,*]

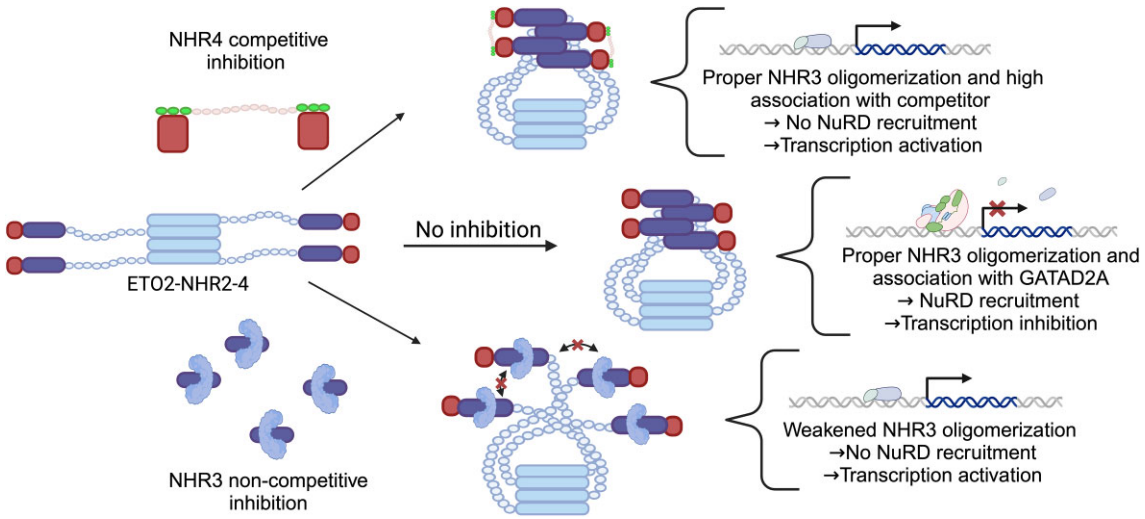
[1]Department of Chemical Biology and Medicinal Chemistry, University of North Carolina, Chapel Hill, NC 27599, United States
[2]Massey Cancer Center, Virginia Commonwealth University, Richmond, VA 23298, United States
[3]Department of Pathology and Laboratory Medicine, University of North Carolina, Chapel Hill, NC 27599, United States
[4]Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, United States
[5]Departments of Internal Medicine and Human and Molecular Genetics, Virginia Commonwealth University, Richmond, VA 23298, United States

[*]To whom correspondence should be addressed. Email: david_willjr@med.unc.edu

Abstract

Over the past 50 years, research has uncovered the co-regulatory proteins and complexes that silence the expression of the γ-globin gene in a developmental stage-specific manner. Recent research expanded the list of these regulatory factors by showing that the eight twenty-one protein 2 (ETO2) helps recruit the nucleosome remodeling and deacetylase (NuRD) complex to the globin locus. Furthermore, ETO2 regulates hematopoietic differentiation and is a potential therapeutic target for acute leukemia. In this work, we identify critical interactions between ETO2 and the GATA Zn finger domain containing the 2A (GATAD2A) component of NuRD. The ETO2 nervy homology region 4 (NHR4) domain interacts with multiple polyproline-leucine motifs within GATAD2A. We demonstrate that oligomerization of the ETO2 nervy homology region 3 (NHR3) enhances its affinity for peptides containing at least two polyproline-leucine motifs. Replacing the native motifs from GATAD2A with a higher-affinity sequence from known-binder *N*-CoR markedly enhances binding affinity, yielding a peptide that disrupts the interaction between ETO2 and target proteins. Enforced peptide expression elevates γ-globin expression levels and induces differentiation of HUDEP-2 and K562 cells. These findings provide insight into ETO2-mediated recruitment of co-regulatory proteins and yield a novel approach for ETO2 inhibition through multivalent binding of the NHR4 domain.

Graphical abstract



Received: September 9, 2024. **Revised:** April 4, 2025. **Editorial Decision:** May 6, 2025. **Accepted:** May 12, 2025

Introduction

The eight twenty-one protein 2 (ETO2) and homologs (ETO and ETO3) play a key role in hematopoietic differentiation [1–5] and, as such, the development of acute myeloid leukemia. They function as co-regulatory proteins by interacting with transcription factors and recruiting chromatin-modifying complexes that predominantly silence transcription of the associated genes. The ETO gene was first identified as a fusion partner in the most common chromosomal translocation in acute myeloid leukemia between chromosomes 8 and 21 (t (8;21)) [6–11], while ETO2 has been identified in less common translocations associated with aggressive forms of leukemia [12–14]. In addition, ETO2 contributes broadly to the differentiation block driving acute leukemia, such that loss of ETO2 induces differentiation of leukemic cell lines [15]

Beyond the relevance of ETO2 in leukemia, recent work by the Dean group identified a novel functional role for ETO2 in silencing the fetal γ-globin gene [3, 16]. They showed that ETO2 localized to the γ-globin promoter and contributed to the recruitment of the nucleosome remodeling and deacetylase (NuRD) complex and silencing of fetal hemoglobin expression during erythroid development. That work was the first to demonstrate that ETO2 directly binds and recruits NuRD for gene silencing; however, the molecular details of this interaction have yet to be characterized. As we and others have demonstrated, the MBD2-NuRD complex plays a critical role in silencing γ-globin expression, and disrupting its formation induces robust fetal hemoglobin expression [17–24]. Notably, inducing fetal hemoglobin expression in adults with β-hemoglobinopathies (sickle cell anemia and β-thalassemia) ameliorates disease symptoms and has been a long-standing therapeutic goal. Hence, we sought to understand how ETO2 binds to the NuRD complex.

Members of the ETO family of proteins contain four conserved nervy homology region domains (NHR1-4) homologous to the Drosophila Nervy protein (Fig. 1). The NHR1 and NHR2 domains interact with transcription factors that recruit ETO to DNA [25–27]. Important for this work, the NHR2 domain was previously shown to form a stable tetrameric antiparallel α-helical bundle (Supplementary Fig. S1A) critical for function. [26, 27] Less is known about the structure of the NHR3 region other than it binds to a small domain from protein kinase A [28] and likely forms a tetramer [29]. Previous research showed that the NHR4 domain (a myeloid, Nervy, and DEAF-1; or MYND domain) binds to small polyproline-leucine (PPPL) motifs from N-CoR and SMRT proteins [30, 31].

The Bushweller group determined the solution structure of ETO-NHR4 bound to a PPPL motif from SMRT [32], demonstrating that the first proline from this motif forms a prolyl-π interaction with a conserved tryptophan, while leucine inserts into a nearby hydrophobic pocket (Supplementary Fig. S1B). Other MYND domain proteins show similar modes of interaction with small peptide motifs [33]. Furthermore, studies by the Miller and Vermeulen groups showed that the zinc finger MYND-type containing eight (Zmynd8) protein recruits NuRD to sites of DNA damage through binding to multiple PPPL motifs within the GATAD2A protein [34, 35]. These previous studies led us to hypothesize that the ETO2 NHR4 domain binds to the same PPPL motifs in GATAD2A and that this interaction contributes to the recruitment of NuRD to the γ-globin locus.

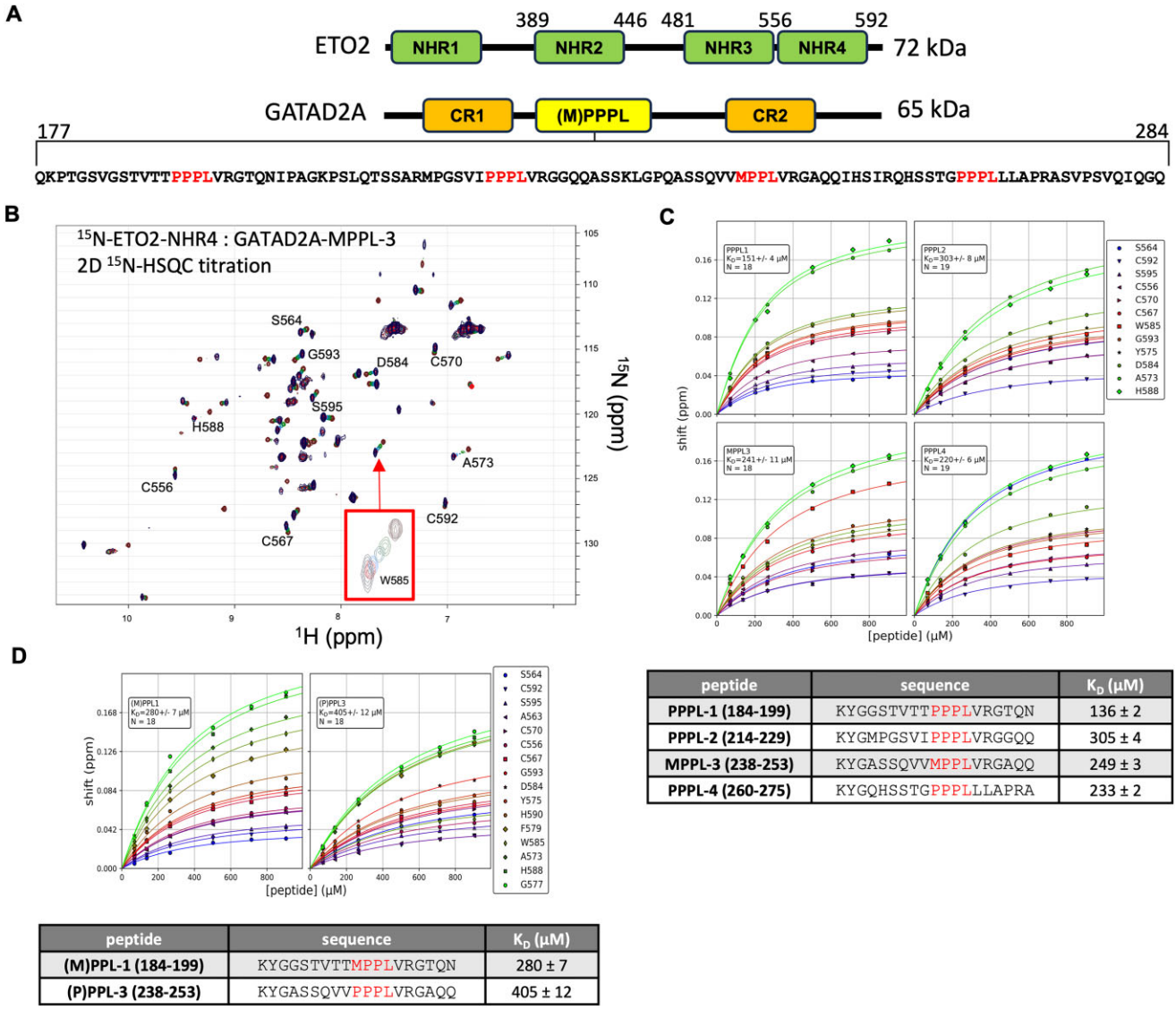
Therefore, to elucidate the molecular details of ETO2-mediated recruitment of GATAD2A-NuRD, we utilized biophysical techniques, cellular analyses, and molecular modeling to study the interaction between ETO2 and GATD2A. We investigated the complex formed between the NHR4 domain and individual PPPL motifs and the contribution of the NHR3 domain to oligomerization and multivalent association. Our findings demonstrate that building a multimeric complex by ETO2 is essential for high-affinity association with target peptides. We show that a high-affinity multivalent peptide blocks binding between ETO2 and native targets, which activates γ-globin expression and induces differentiation of leukemia cells, a proven approach for treating acute leukemia [15, 36]. Therefore, this peptide suggests utilizing multivalency to selectively and potently target ETO proteins and possibly other MYND domain proteins as a novel molecular probe.

Materials and methods

Expression and purification of proteins

We cloned the NHR-3–4 and NHR4 domains of human ETO2 (residues 480–597 and 551–597) into a modified pET28a vector containing *N*-terminal hexa-histidine (6xHis) and maltose binding protein (MBP) tags with an intervening tobacco etch virus (TEV) protease cleavage site. We transformed the vectors into Rosetta2 (DE3) *E. coli* (Invitrogen), incubated at 37°C in Lysogeny Broth (LB) media with kanamycin antibiotic selection (50 μg/mL), and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at an OD_{600} of ∼0.8. After 3 h, bacteria were harvested by centrifugation and stored at −20°C. The bacterial pellets were resuspended in buffer containing 20 mM Tris pH 8.0, 200 mM NaCl, 1 mM dithiothreitol (DTT), a Pierce protease inhibitor tablet (Thermo Scientific), or 1 mM phenylmethylsulphonyl fluoride (PMSF), and 50 μM $ZnSO_4$. The bacterial suspension was lysed by sonification and centrifuged at 12 500 × g. The supernatant was purified over 2 × 6 mL amylose resin (New England Biolabs) and eluted by TEV protease digestion. The eluted protein was purified via nickel affinity chromatography (HisTrap Fast Flow, Cytiva) followed by size-exclusion chromatography (Superdex-75 Increase 10/300, GE Healthcare, or HiLoad 26/600 Superdex 75 pg, GE Healthcare) in 25 mM Bis-Tris pH 7.2, 100 mM NaCl, 1 mM DTT, and 50 μM $ZnSO_4$.

For (M)PPPL peptides containing more than one binding site, we cloned residues 177–284 ((M)PPPLx4), 177–235 (PPPLx2A), 205–261 ((M)PPPLx2B), and 229–284 ((M)PPPLx2C) of GATAD2A into a modified pET32a vector containing *N*-terminal thioredoxin and 6xHis tags and a TEV protease site and a *C*-terminal Twin-Strep-tag® [37] following a short linker (GGSGGSSA). For N-CoR-modified peptides, we replaced the (M)PPPL motifs of GATD2A with residues 1031–1041 of N-CoR. Vectors were transformed into Rosetta2 (DE3), grown in LB + ampicillin media, and induced following the same procedures as for the ETO2 domain proteins. The pellet was resuspended in 20 mM TRIS pH 8.0, 1M NaCl, 1 mM DTT, and a Pierce protease inhibitor tablet (Thermo Scientific) or 1 mM PMSF. Following sonification and centrifugation, the sample was purified with nickel affinity chromatography and eluted in 20 mM Tris pH 8.0, 300 mM NaCl, 300 mM imidazole, and 1 mM DTT. The elution was digested overnight with TEV



peptide	sequence	K_D (µM)
PPPL-1 (184-199)	KYGGSTVTTPPPLVRGTQN	136 ± 2
PPPL-2 (214-229)	KYGMPGSVIPPPLVRGGQQ	305 ± 4
MPPL-3 (238-253)	KYGASSQVVMPPLVRGAQQ	249 ± 3
PPPL-4 (260-275)	KYGQHSSTGPPPLLLAPRA	233 ± 2

peptide	sequence	K_D (µM)
(M)PPL-1 (184-199)	KYGGSTVTTMPPLVRGTQN	280 ± 7
(P)PPL-3 (238-253)	KYGASSQVVPPPLVRGAQQ	405 ± 12

Figure 1. ^{15}N-HSQC chemical shift perturbation experiments show that the individual (M) PPPL motifs weakly bind to ETO2-NHR4. (**A**) ETO2 contains four conserved Nervy Homology Regions (NHR1, NHR2, NHR3, and NHR4). GATAD2A contains two highly Conserved Regions (CR1 and CR2), with an intervening region containing four neighboring polyproline-leucine motifs ((**M**)PPPL), as shown in the amino acid sequence of this region. (**B**) 2D ^{15}N-HSQC spectra of ^{15}N-ETO2-NHR4 show that titration of GATAD2A-MPPPL-3 induces chemical shift changes throughout the spectrum, which is highlighted for W585 (inset). (**C**) The chemical shift perturbation data for 12 resonances were fit simultaneously to determine dissociation constants (K_D) for PPPL-1, PPPL-2, MPPL-3, and PPPL-4. (**D**) Chemical shift perturbation data for (**M**)PPPL-1 and (**P**)PPPL-3 show that interchanging the first proline or methionine of the binding motifs with methionine or proline, respectively, reduces the binding affinity for each.

and then purified over a Strep-Tactin®XT 4Flow® FPLC column (IBA-Lifesciences) and eluted with 1x Buffer BXT (IBA-Lifesciences).

For the purification of the ETO2-NHR3 coiled-coil region, we cloned residues 480–537 of ETO2 into the same modified pET32a vector containing *N*-terminal thioredoxin and 6-His tags and a TEV protease restriction site. We expressed and purified the domains with nickel affinity chromatography as described for the (M)PPPL peptides. The sample was digested with TEV protease and dialyzed overnight against 20 mM Tris pH 8.0, 150 mM NaCl, 1 mM DTT, then sequentially purified by nickel and size exclusion chromatography.

Residues 1–45 of cAMP-dependent kinase type II-alpha regulatory subunit (PKA (RIIα)) were cloned into the same modified pET32a vector and purified with the same protocol as the ETO2-NHR3 coiled-coil region.

Peptide synthesis

For NMR titration analysis and ITC, we synthesized peptides containing a single polyproline-leucine motif via solid-phase with a CEM Liberty Blue peptide synthesizer on Rink Amide ProTide resin (LL) using Fmoc-protected amino acids. Peptides were purified by reverse-phase HPLC and verified by ESI-LCMS.

NMR spectroscopy

For structure determination, we cloned the MPPL-3 peptide sequence (GATAD2A residues 238–253) as an *N*-terminal fusion with the ETO2-NHR4 domain (residues 551–597) separated by a short linker comprised of the amino acid sequence "KNGRGS" into the same modified pET28a vector with an *N*-terminal MBP tag described above. Following established

techniques, we isotopically labeled ETO2-NHR4 and ETO2-NHR4-MPPLsc proteins by expression in M9 media and purified as described above.

We prepared all NMR samples in 25 mM BisTris, pH 6.8, 100 mM NaCl, 1 mM DTT, and 25 μM $ZnSO_4$, with 5–10% ^2H_2O. We collected standard 2D and 3D NMR spectra for assignments, binding analysis, and distance restraints on a Bruker Avance III HD 850 MHz four-channel spectrometer equipped with a TCI H-C/N D 5 mm cryoprobe at 25°C, processed with NMRPipe [38], and analyzed with CCPNmr Analysis versions 2 and 3 [39]. We collected residual dipolar couplings for samples aligned in 5% C12E5 poly (ethylene glycol)/*n*-hexanol ($r = 0.85$) (Sigma) at 22 °C [40].

Gel filtration analysis

We loaded ETO2 at various concentrations (25, 100, 500 μM) and compared the elution position with Zmynd8 (25 μM) on a Superdex-75 10/300 size exclusion chromatography column (Cytiva). To assess the multimeric complex formation of ETO2 and target peptides, ETO2-NHR3-4 (60 μM) was injected with ligand (N-CoRx2A, N-CoRx4) at a 2:1 or 4:1 molar ratios (30 and 15 μM). We compared the elution profile to that of the isolated ETO2-NHR3-4 and the isolated ligands, N-CoRx2A or N-CoRx4. We performed all gel filtration experiments under identical buffer conditions (25 mM Bis-Tris pH 7.2, 100 mM NaCl, 1 mM DTT, and 50 μM $ZnSO_4$).

NanoBRET™ endpoint detection in HEK293T/17 cells

We cloned the different ETO2 domain and GATD2A polyproline-leucine peptides into the EcoRI sites of the HaloTag® (pHTC) and NanoLuc® (pNLF1-C or pNFL1-N) vectors (Promega), respectively. For the initial experiments with PPPLx4, we included the coiled-coil domain from GATAD2A (CR1) fused to the coiled-coil domain from MBD2 as a single chain, as we have described previously [41–43]. We cultured HEK293T/17 cells in DMEM media supplemented with 10% FBS at 37°C and 5% CO_2. We transfected cells at ~80% confluency in a six-well (or 12-well) cell-culture treated plates with 2 μg (or 1) and 0.2 μg (or 0.1) of HaloTag® and NanoLuc® fusion proteins using 6 μL (or 3) of Fugene HD Transfection Reagent (Promega) following the manufacturer's protocol. After 18–24 h, we washed and resuspended cells in Opti-MEM I Reduced Serum medium, no phenol red, supplemented with 4% FBS, and replated onto 96-well plates. We treated the cells with HaloTag® NanoBRET™ 618 Ligand (Promega) or DMSO for 4–20 h. After the addition of NanoBRET™ NanoGlo® substrate (Promega), we measured donor (460 nm) and acceptor (618 nm) emission on a CLARIOStar Plus Microplate Reader. The resulting BRET signal was calculated as the ratio of acceptor to donor emission (x1000) and corrected by subtracting the no-ligand (DMSO) control. All experiments were performed in quadruplicate and repeated 2–3 times.

Isothermal titration calorimetry (ITC)

All proteins for ITC analysis were extensively exchanged into a buffer containing 25 mM Bis-Tris, pH 7.2, 100 mM NaCl, 1 mM DTT, and 50 μM $ZnSO_4$. The polyproline-leucine peptides, N-CoRx4 (125 μM), (M)PPPLx4 (125 μM), N-CoRx2A (250 μM), or PPPLx2A (250 μM) were injected into the cell containing ETO2 domains, ETO2-NHR4 or ETO2-NHR3-4 (50 μM), for a total of 20 injections (2 μL, 180 s between each injection) at 25°C on either a MicroCal AutoiTC200 or Malvern PEAQ-ITC Fully-Automated Isothermal Titration Calorimeter. For non-competitive inhibition analysis, N-CoRx2A and PPPLx2A (250 μM) were injected into a 2:1 molar ratio of PKA (RII)α (100 μM) and ETO2-NHR3-4 (50 μM). The resulting data were fit using the PEAQ-ITC analysis software.

AlphaFold model of the ETO2-NHR3-4 tetramer

We generated models of the ETO2-NHR3-4 tetramer with and without peptide sequences containing multiple binding sites with AlphaFold-multimer version 2.3.0 [44, 45]. The calculation was performed on a local computer cluster utilizing an A100 40 GB GPU (NVIDIA). All structure analyses and figures were generated with UCSF ChimeraX 1.7.1 [46] with the PICKLUSTER plug-in [47].

ETO2-NHR3 tetramer stabilization

The coordinates of the ETO2-NHR3 domain (residues 480–553) were extracted from the AF2m model and evaluated using the stabilization protocol (Stabilize Proteins: PM) on the Rosetta Online Server that Includes Everyone (ROSIE2) [48, 49]. We introduced symmetric mutations at each position (residues 481–553) and ranked them according to the predicted change in relative energy. We visualized each mutation on the native structure, prioritizing those that improved the packing of the hydrophobic core at the tetramer interface. We repeated the analysis, starting with the coordinates of the three top-ranked AF2m models, and based on these results, we introduced the Ala501Leu in isolation and combined it with Val497Ile. We generated a model of the ETO2-NHR3-4 (A501L) and ETO2-NHR3-4 (V497I/A501L) tetramer with AF2m, as described for the wild-type protein.

Circular dichroism

We exchanged the purified ETO2 coiled-coil domains into a buffer containing 10 mM KH_2PO_4, pH 7.5, and 100 mM NH_4SO_4 at a final concentration of 0.15 mg/mL (22 μM). We collected circular dichroism spectra on a Chirascan™-Plus CD Spectrometer (AppliedPhotophysics) at 20°C, scanning from 190 to 260 nm at 0.5 nm intervals and 0.4 nm/s. We normalized the resulting spectra and calculated the helical content based on molar ellipticity at 222 nm [50].

Size exclusion chromatography with multi-angle light scattering (SEC-MALS)

We performed SEC-MALS experiments in a buffer containing 25 mM Bis-Tris pH 7.2, 100 mM NaCl, 1 mM DTT, 50 mM $ZnSO_4$, and 200 mg/L NaN_3 at 25°C on instrumentation available in the UNC Macromolecular Interactions Core Facility. Samples (100 μL) of wild-type or mutant ETO2-NHR3-4 (120 μM) with N-CoRx2A (60 μM) were injected onto a Superdex 200 Increase 10/300 GL (GE Healthcare) using an AKTA Pure (GE Healthcare) in line with a DAWN HELEOS II MALS detector (Wyatt Technology). Data were analyzed with the ASTRA 6 software (Wyatt Technology) utilizing a Zimm light scattering model with a dn/dc value of 0.1850 and an albumin control.

Cell culture and differentiation

HUDEP-2 cells were maintained and expanded in StemSpan SFEM II supplemented with 50 ng/mL of stem cell factor, 3 IU/mL erythropoietin (EPO), 1 μM of dexamethasone (DEX), 1 μg/mL of doxycycline (DOX), 1% L-glutamine, and 2% penicillin/streptomycin (expansion medium). To induce erythroid differentiation, HUDEP-2 cells were cultured for 3 days in IMDM supplemented with 3 IU/mL EPO, 1 μg/mL DOX, 10 μg/mL human recombinant insulin, 3 IU/mL heparin, 0.5 mg/mL of HTF, 5% human AB serum, 1% penicillin/streptomycin, and 1% L-glutamine (differentiation medium). K562 cells were cultured in RPMI-1640 supplemented with 10% fetal bovine serum (FBS), 1% sodium pyruvate, 1% L-glutamine and 1% penicillin/streptomycin. The cells were cultured at 37°C in 5% CO_2.

Lentivirus vector preparation and infection

Around 5×10^6 293T cells were seeded into 10 cm culture dishes with 10 mL DMEM supplemented with 10% FBS, 1% HEPES, 1% non-essential amino acid (NEAA), 1% penicillin/streptomycin, and 1% L-glutamine. After 24 h, the cells were given fresh media and co-transfected with 8 μg of either lentiviral expression plasmid containing N-CoRx2A with a nuclear localization signal or empty vector together with 6 μg psPAX2 and 4 μg pMD2.G using polyethyleneimine (PEI). 12–16 h post-transfection, the media was replaced by 5mL DMEM supplemented with 5% FBS, 1% HEPES, and 1% NEAA. Viral supernatants were collected 48 h post-transfection, filtered through 0.45 μm micropore filters, and used while still fresh to infect HUDEP-2 and K562 cells.

StrepTag immunoprecipitation and electrophoresis

Around 3×10^7 K562 cells were collected, subjected to nuclear protein extraction [51], and treated with MNase digestion [18]. Following centrifugation, the cleared cell lysate was immunoprecipitated with MagStrep "type3" XT beads using the manufacturer's protocols. Samples were then boiled in 1 × sodium dodecyl sulphate (SDS) loading buffer for 5 min to denature all proteins and separated by 10% SDS-polyacrylamide gelelectrophoresis.

Antibodies and reagents

Flag antibody (F1804-200UG) was purchased from Sigma-Aldrich. CHD4 antibody (12 011 s) was purchased from Cell Signaling Technology. ETO (sc-134335), ETO2 (sc-166058), and Vinculin (sc-73614) antibodies used in this study were purchased from Santa Cruz Biotechnology. The Strep-Tactin magnetic kit (MagStrep "type3" XT beads were obtained from IBA) was used for Strep immunoprecipitation.

Results

ETO2-NHR4 weakly binds the individual (M)PPPL motifs of GATAD2A-NuRD

Previous work showed that the Zn-finger MYND-type containing 8 (Zmynd8) protein binds to three PPPL motifs within GATAD2A [34, 35]. Given that ETO2-NHR4 functions in co-factor recruitment [32, 52], we hypothesized it directly binds these same motifs. In addition, we noted a fourth potential binding site with a slightly modified motif (MPPL) in the same region of GATAD2A (Fig. 1A). We measured binding to these four sites by NMR spectroscopy. We synthesized 18-residue peptides containing each motif and followed binding with ^{1}H-^{15}N heteronuclear single quantum correlation (HSQC) spectra of ^{15}N-labeled ETO2-NHR4 (amino acids 551–597) (Fig. 1B). This analysis revealed that the ETO2-NHR4 domain binds each of the four motifs within GATD2A with dissociation constants (K_D) ranging from 136 to 305 μM (Fig. 1C and D). The chemical shift changes upon binding are similar among all four peptides to that reported for the ETO-NHR4-SMRT complex, indicating they bind analogously. Notably, the third motif within the primary sequence of GATAD2A (MPPL-3) has a similar affinity ($K_D = 249 \pm 3$ μM), falling in the middle of the range, despite the presence of a methionine instead of proline as the first residue (Fig. 1C).

The previous structural analysis of ETO-NHR4 binding to SMRT showed that the first proline in the PPPL motif forms a prolyl–π interaction with a conserved tryptophan in the MYND domain (Supplementary Fig. S1B) [32]. Consequently, we questioned whether replacing this proline with a methionine was detrimental to the binding affinity. We synthesized two modified binding sites to test this possibility by replacing the methionine in MPPL-3 with proline and the first proline with methionine in PPPL-1 (the highest affinity binding site in GATAD2A). The two modified peptides showed a small reduction in binding affinity for ETO2-NHR4, underscoring the importance of the surrounding residues in the native sequences (Fig. 1D). Therefore, in the context of the MPPL-3 peptide sequence, methionine is the preferred residue in the first position of the motif, whereas, in the context of the PPPL-1 sequence, proline is preferred.

Solution structure of ETO2-NHR4:MPPL-3 complex

To further explore the role of methionine in this binding site, we determined the structure of ETO2-NHR4 bound to MPPL-3. Given the weak binding affinity, we followed the same approach as Liu *et al.* [32] by expressing MPPL-3 and the NHR4 domain as a single polypeptide separated by a short linker (6 amino acids) (Fig. 2A). Comparing the 2D-^{15}N-HSQC spectrum of the ETO2-NHR4-MPPL-3 single-chain (MPPLsc) with the MPPL-3 peptide titration experiment confirms that the linker stabilizes the bound state, such that the chemical shift changes exceed that of the highest peptide-protein ratio and does not alter the mode of interaction (Fig. 2B). The solution structure (Fig. 2C and D and Supplementary Fig. S2) was well determined, with a backbone RMSD of 0.6 Å for structured residues. As expected, the structure of the ETO2-NHR4 domain is the same as ETO-NHR4 [32], with a backbone root mean square deviation of 1.0 Å between the two MYND domains while in complex with peptide (PDBID: 2odd). The overall fold comprises two zinc-coordinating centers formed by Cys 556, 559, 576, and 580 and Cys 567, 570, 592, and His 588. These two zinc-binding sites involve both ends of the only alpha-helix, inducing it to bend around Tyr 575 (Fig. 2D), which occupies the core of the domain. These features are common to all MYND domain structures solved to date [32, 33, 53–56].

As anticipated, the MPPL-3 peptide binds in the same orientation as the SMRT motif, with methionine packing against Trp585 and the leucine sidechain buried in a deep pocket between Gln581, Glu565, and Tyr575 (Fig. 2C). The two central prolines bridge these two interaction sites, adopting a polyproline type II helical backbone structure that arches over a ridge

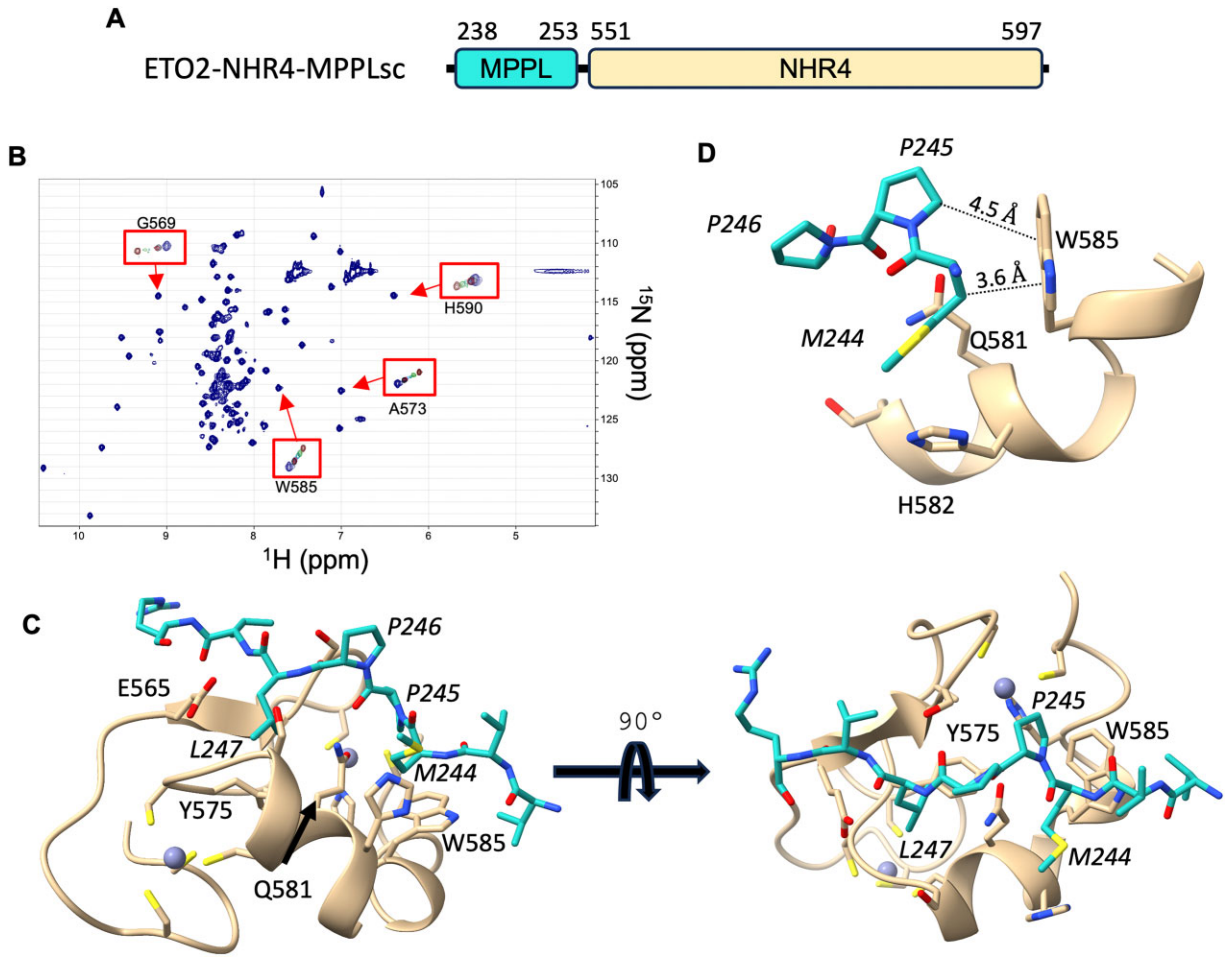


Figure 2. Solution structure of ETO2-NHR4 bound to MPPPL-3. (**A**) A diagram depicts the ETO2-NHR4-MPPLsc fusion used for this structural analysis. (**B**) 2D ^{15}N-HSQC spectrum of ^{13}C,^{15}N-ETO2-NHR4-MPPLsc with insets showing titration spectra for G569, W585, A573, and H590 resonances (dark blue represents ^{15}N-ETO2-NHR4-MPPLsc, and maroon, green, light blue, red, and gray represent titration of MPPL-3 into ^{15}N-ETO2-NHR4). (**C**) The solution structure of ETO2-NHR4-MPPLsc shows the MPPL-3 peptide (cyan) adopts a polyproline-II helical structure that binds in a groove on the surface of the NHR4 domain, with the first methionine (M244) interacting with the conserved tryptophan (W585) and the leucine (L247) binding in a deep pocket, the base of which is formed by a tyrosine (Y575). (**D**) An expanded view of the binding interface shows that the Cα and Cβ atoms of M244 and the Cδ and Cγ atoms of P245 pack against the pyrrole and benzene portions of W585, respectively.

formed by the sidechain of Gln581. This latter sidechain is appropriately positioned to make a hydrogen bond with the backbone carbonyl of the first proline in the MPPL motif. Interestingly, the methionine points away from the interface such that the Cα and Cβ atoms pack against the pyrrole portion of the Trp585 indole, and the subsequent proline Cδ and Cγ atoms pack against the benzene portion (Fig. 2D). This orientation indicates that the interaction is predominantly hydrophobic and argues against forming a sulfur–π interaction with Trp585 [57].

Multivalency drives higher affinity association of the NHR4 and NHR3-4 domains of ETO2 to GATAD2A

The NHR3 region of ETO proteins immediately precedes the NHR4 domain and is predicted to form a coiled-coil domain. We hypothesized that oligomerization of NHR3 would bring multiple NHR4s in proximity, allowing them to associate with contiguous binding sites in GATAD2A simultaneously. Therefore, we predicted that constructs including both NHR3 and NHR4 domains would show a higher apparent binding affinity for peptides encompassing multiple (M)PPPL motifs from GATAD2A. To investigate this possibility, we utilized an intracellular bioluminescence resonance energy transfer proximity-based assay, NanoBRET™, to measure binding between different ETO2 and GATAD2A domains. We designed HaloTag® and NanoLuc® constructs containing various combinations of (M)PPPL motifs and ETO2 domains, respectively (Fig. 3A).

First, we tested whether the NHR2 and NHR3 contributed to stable association. As shown in Fig. 3B, the BRET ratio increases with the addition of NHR3 and NHR2 domains. Importantly, we used a C-terminal HaloTag® and truncated ETO2 from the *N*-terminus to maintain the same relative position of the BRET acceptor. These results indicate that NHR2 and NHR3 domains help stabilize binding to the (M)PPPLx4 region (ETO2-NHR2-4 > ETO2-NHR3-4 > ETO2-NHR4). Mutating Trp585 disrupts the interaction for all constructs, confirming that binding involves the NHR4–PPPL interface. Next, we tested whether the number of (M)PPPL motifs impacted the interaction with ETO2-NHR2-4. Varying the number of (M)PPPL motifs binding to ETO2-NHR2-4 shows increasing BRET ratios as the number of motifs increases

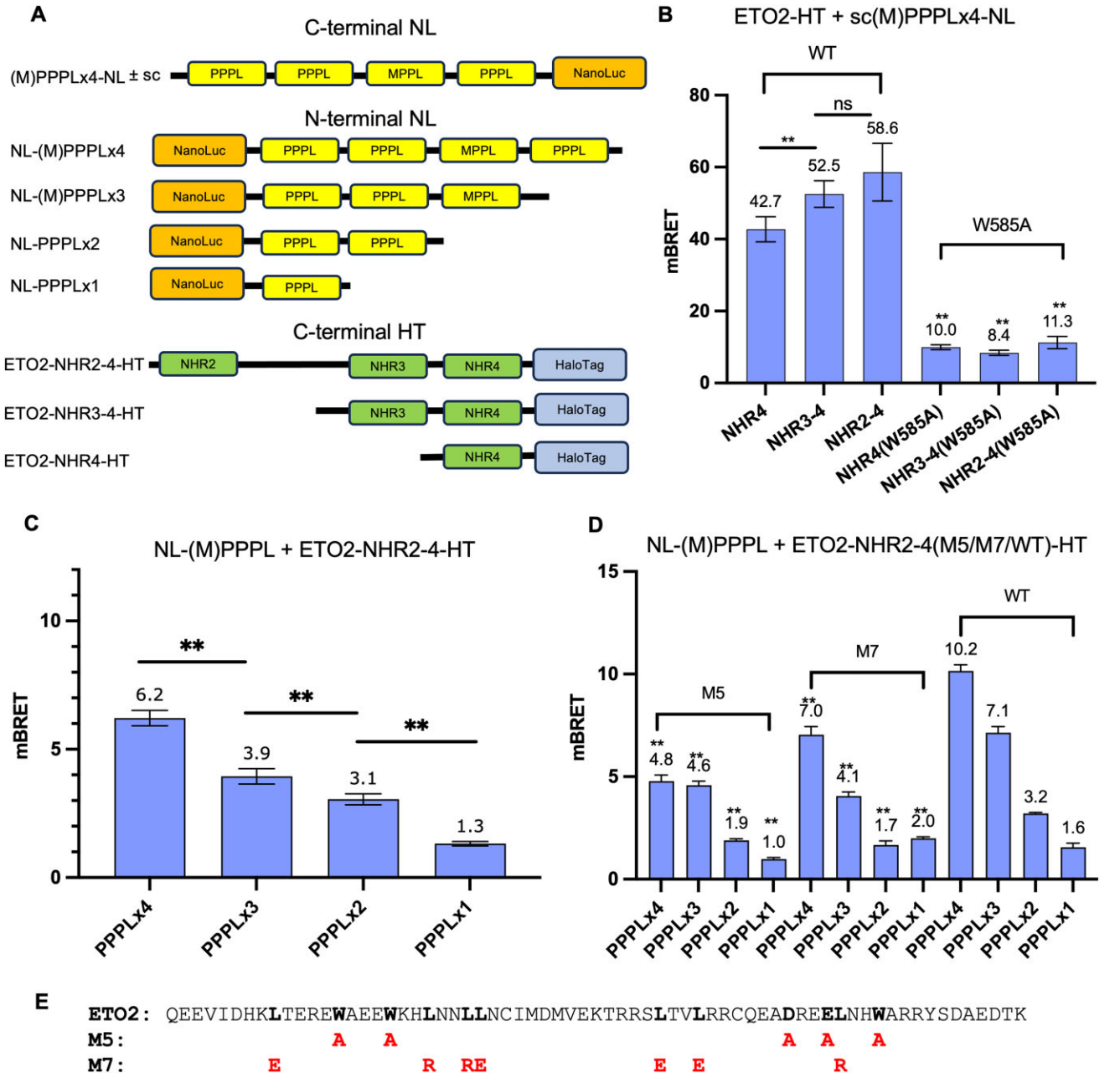


Figure 3. In-cell binding analysis between ETO2 and the GATAD2A (M) PPPL region. (**A**) A diagram depicts the different NanoLuc (NL) and HaloTag (HT) constructs for binding analysis by bioluminescence resonance energy transfer (NanoBRET). Six different (**M**)PPPL constructs were designed, including two with the NL donor tag on the *C*-termini (one including the neighboring CR1 domain as a coiled-coil single chain sequence on the *N*-terminus, see Materials and methods). The remaining four (**M**)PPPL constructs contain the NL tag on the *N*-terminus to allow for *C*-terminal truncation while maintaining the same relative position of the NL tag. Three ETO2 *C*-terminal HT constructs were designed to maintain the same relative position of the HT acceptor tag. (**B**) The interaction between ETO2-NHR4 and sc (**M**)PPPLx4 yields a strong BRET intensity, which progressively increases with the inclusion of the NHR3 and NHR2 domains. Mutating the key tryptophan residue that interacts with the first proline of the PPPL motif (M585A) markedly decreases the BRET signal with all ETO2-HT constructs. (**C**) Truncating the PPPLx4 progressively reduces the observed BRET signal. (**D**) Introducing mutations to the ETO2-NHR2 tetramerization domain that disrupt tetramerization (M7) or dimerization (M5) decreases the BRET signal intensity (*P* values reflect a comparison with the equivalent wild-type interaction). (**E**) A diagram shows the M5 and M7 mutations in the ETO2-NHR2 tetramerization domain. All measurements were performed in quadruplicate and repeated at least twice to confirm the results. *P*-values (<0.05=* and <0.01=**) were determined by the Student's T-test with Welch's correction.

(Fig. 3C). The results indicate that (M)PPPLx4 binds most stably to ETO2-NHR2-4.

The NHR2 domain of ETO proteins forms a stable antiparallel tetramer comprising a dimer of dimers (Supplementary Fig. S1A). Previous work by Wichmann *et al.* [58] identified mutations that selectively disrupted the tetramerization interface without disrupting dimerization (M5) and mutations that blocked dimerization to yield monomeric species (M7) (Fig. 3E). Hence, to elucidate the effect of either dimerization or tetramerization of the NHR2 domain on binding, we introduced these mutations and measured associations with GATAD2A by NanoBRET™. Both sets of mutations reduced the BRET ratios similarly (Fig. 3D), reducing the interaction with each of the (M)PPPLx1-4 constructs.

In addition to intracellular assays, we measured *in vitro* binding affinities by isothermal titration calorimetry (ITC). Different ETO2 domains (NHR4 or NHR3-4) and peptides containing either four (M)PPPL motifs ((M)PPPLx4) or three pairwise combinations of two consecutive motifs ((M)PPPLx2A/B/C) were expressed in bacteria and purified (Fig. 4A and B). The results revealed that the ETO2-NHR4 binds (M)PPPLx4 with a K_D of 30 ± 5 μM, such that including all four motifs increases the binding affinity up to 10-fold (Fig. 4C and Supplementary Fig. S3A) compared to the individual binding sites. Including the NHR3 domain (ETO2-NHR3-4) increases the binding affinity an additional 10-fold (K_D = 3 ± 1 μM) compared to the NHR4 domain alone (Fig. 4C and Supplementary Fig. S3B). Based on this observation, we tested whether stable interaction required all four binding sites. Peptides containing the first, middle, and last two motifs (PPPLx2A, (M)PPPLx2B, and (M)PPPLx2C) were purified, and binding affinities with ETO2-NHR3-4 were measured. The results show that the first two motifs (PPPLx2A) were sufficient to increase the binding affinity to a similar level as all four (K_D = 6 ± 5 μM), while the second and third pairs did not bind as tightly (Fig. 4D and Supplementary Fig. S4). These results suggest that the first two motifs have the appropriate spacing and linker amino acid sequence for optimal multivalent association with ETO2-NHR3-4.

Structural analysis of ETO2-NHR3 oligomerization

The zinc finger MYND domain-containing (Zmynd) protein family members frequently contain coiled-coil regions immediately preceding the MYND domain. For example, previous structural analysis of Zmynd11 (BS69) showed that its coiled-coil domain forms a stable parallel dimer that contributes to binding a multivalent ligand [33]. In addition, a crystal structure of Zmynd8 (Supplementary Fig. S1C, PDBID: 5mq4) shows its coiled-coil domain forms a parallel tetramer. Therefore, we hypothesized that the ETO2-NHR3 domain would form a tetramer that promotes multivalent association with GATAD2A. We first performed gel filtration analyses to test this possibility. As Fig. 5A shows, the purified ETO2-NHR3-4 protein elutes at progressively earlier positions as the injected concentration ranges from 25 to 500 μM. For comparison, Zmynd8 (25 μM) elutes as a single narrow peak at a position comparable to the highest concentration of ETO2-NHR3-4. These results indicate that ETO2-NHR3-4 weakly oligomerizes, likely forming a homo-tetramer.

To gain insight into the formation of tetrameric ETO2-NHR3-4, we used AlphaFold2-multimer (AF2m) to predict its structure. AF2m consistently generates models with an antiparallel arrangement for ETO2-NHR3-4 (Fig. 5B). The predicted local distance difference test (pLDTT) scores of the top models are in the high confidence range (70–90). Likewise, the position alignment error (PAE) plot highlights the well-defined interface between the coiled-coil domains.

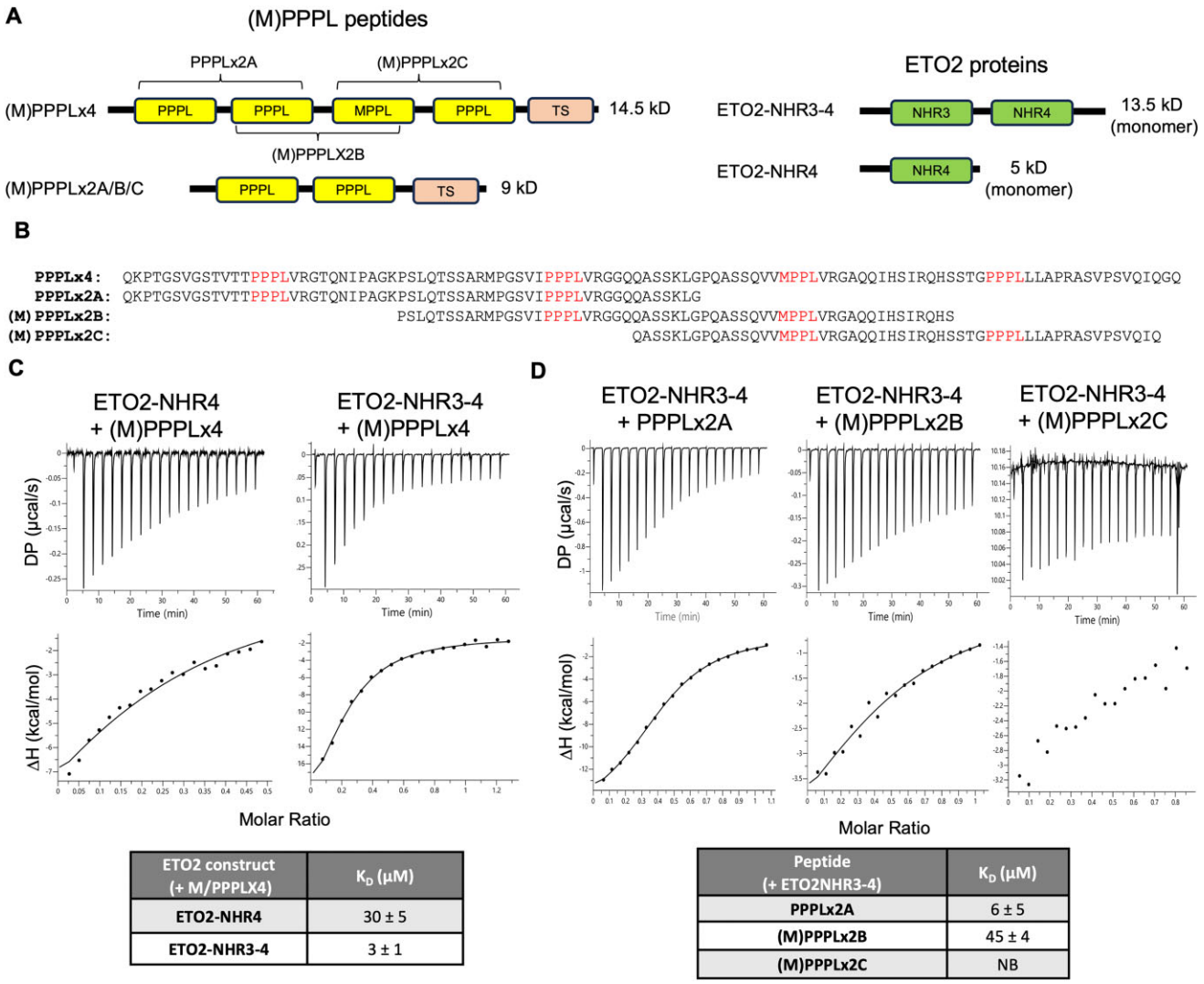
In the best-ranked models, ETO2-NHR3 forms two antiparallel coiled-coil dimers, with hydrophobic residues packed at the interface (Fig. 5B), yielding a dimer of dimers. The tetrameric helical bundle involves the first ∼45 residues of the NHR3 domain (residues 481–526), while the remaining ∼27 residues (527–554) form a parallel coiled-coil dimer. The antiparallel dimer involves an extensive and well-defined interface (Fig. 5B, highlighted in red) with an average pLDDT score of 75 and an average PAE score of 5.1 Å. The two antiparallel dimers interact through a second extensive antiparallel interface (highlighted in blue) with an average pLDDT score of 74 and an average PAE of 3.7 Å. The helices that extend beyond the end of the tetramer then loosely interact as parallel dimers (highlighted in orange) that are not as well defined with an average pLDDT score of 64 and average PAE of 8.9 Å.

The antiparallel tetrameric structure results in two NHR4 domains on opposite ends of the helical bundle, such that each pair can interact with two (M)PPPL motifs in GATAD2A. Therefore, we generated a model of a complex between the ETO2-NHR3-4 tetramer with one copy of each of the PPPLx2A and (M)PPPLx2C peptides. As expected, the two peptides bind on opposite ends of the tetramer, with each (M)PPPL binding site interacting with a single ETO2-NHR4 domain (Fig. 5C). The ETO2-NHR3 region retains the same antiparallel structure with similar confidence scores. In the complex, the ETO2-NHR4 domains consistently adopt a more extended structure, which allows them to interact with the bivalent peptides. However, the confidence scores for the peptides are much lower (pLDTT 50–70) than for the ETO2 domains. In addition, the separation between NHR4 domains differs among models, reflecting a lack of direct interaction between them (Supplementary Fig. S5).

The AF2m model provides insight into the multivalent association with GATAD2A. First, the antiparallel tetramer presents a pair of ETO2-NHR4 domains on each end of a long helical bundle. This arrangement helps explain why PPPLx2A binds with nearly identical affinity as (M)PPPLx4, suggesting that the separation between the ends prevents simultaneous interaction with all four binding sites in a single polypeptide. Furthermore, the relative binding affinities between PPPLx2A, B, and C suggest that either the length or composition of the linker separating the two sites contributes to the multivalent binding. To test these ideas, we replaced the four different binding motifs with the same amino acid sequence to directly compare the impact of the linkers and the number of binding sites on binding affinity.

The binding motif from N-CoR stabilizes tetramer formation and promotes higher affinity association

In the seminal work identifying the MYND domain binding motif, Liu *et al.* showed that a peptide from the N-CoR protein bound to ETO-NHR4 with a K_D of 31 μM [32], which is 5–10 fold tighter than any of the GATAD2A binding sites. Therefore, we tested whether substituting the N-CoR peptide sequence for each would increase multivalent



ETO2 construct (+ M/PPPLX4)	K_D (µM)
ETO2-NHR4	30 ± 5
ETO2-NHR3-4	3 ± 1

Peptide (+ ETO2NHR3-4)	K_D (µM)
PPPLx2A	6 ± 5
(M)PPPLx2B	45 ± 4
(M)PPPLx2C	NB

Figure 4. *In vitro* binding affinity analysis of the multivalent interaction between ETO2 and GATAD2A. (**A**) A diagram depicts the various (**M**)PPPL peptides and ETO2 protein constructs for binding analysis. (**B**) The peptide amino acid sequences are shown for PPPLx4, PPPLx2A, (**M**)PPPLx2B, and (**M**)PPPLx2C. (**C**)The ITC binding isotherms and associated fit to a one-site binding model are shown for (**M**)PPPLx4 (125 µM) titrated into ETO2-NHR4 or ETO2-NHR3-4 (50 µM) and (**D**) PPPLx2A, (**M**)PPPx2B, and (**M**)PPPLx2C peptides (250 µM) titrated into ETO2-NHR3-4 (50 µM). The reported dissociation constants (K_D) are the average and standard deviation for 2–3 repeated measurements (see Supplemental Figs S3 and S4).

binding affinity. We replaced 11 residues at each binding site, comprising the (M)PPPL motifs and surrounding amino acids, with 11 residues comprising the higher-affinity peptide from N-CoR, generating N-CoRx4 (Fig. 6A). This modified peptide increased binding affinity for ETO2-NHR3-4 nearly 30-fold (K_D = 110 ± 2 nM, Fig. 6B and Supplementary Fig. S6) compared to (M)PPPLx4 (K_D = 3 ± 1 µM). As with the native sequence, comparing the three pairwise bivalent constructs shows that N-CoRx2A binds with an affinity (K_D = 118 ± 2 nM) similar to that of N-CoRx4 and more tightly than N-CoRx2B (K_D = 2900 ± 300 nM) and N-CoRx2C (K_D = 4000 ± 500 nM).

Having established a high-affinity multivalent ligand with identical binding sites, we compared the N-CoRx4 and N-CoRx2A complexes by gel filtration analysis. As shown in Fig. 6C, the N-CoRx2A:ETO2-NHR3-4 complex (1:2 molar ratio) elutes much earlier than ETO2-NHR3-4 alone, consistent with stable complex formation. Adding excess N-CoRx2A (1:1 molar ratio) does not shift the position of the complex but yields an additional peak consistent with free N-CoRx2A. These results support the expected 1:2 stoichiometry between ETO2-NHR3-4 and N-CoRx2A. In contrast, the N-CoRx4:ETO2-NHR3-4 complex (1:4 molar ratio) elutes even earlier than N-CoRx2A:ETO2-NHR3-4 (Fig. 6D). Notably, adding excess N-CoRx4 (1:2 molar ratio) delays the major peak to a position similar to N-CoRx2:ETO2-NHR3-4 without evidence of excess free peptide. These results indicate that when N-CoRx4 is limiting, the multivalent ligand crosslinks multiple ETO2-NHR3-4 tetramers such that the complex elutes earlier. When in excess, though, ETO2-NHR3-4 preferentially associates with the first two binding sites (equivalent to N-CoRx2A), such that two peptides bind to each tetramer without crosslinking and the complex elutes at a later position, comparable to the complex formed with the bivalent peptide, N-CoRx2A (see Supplementary Fig. S7).

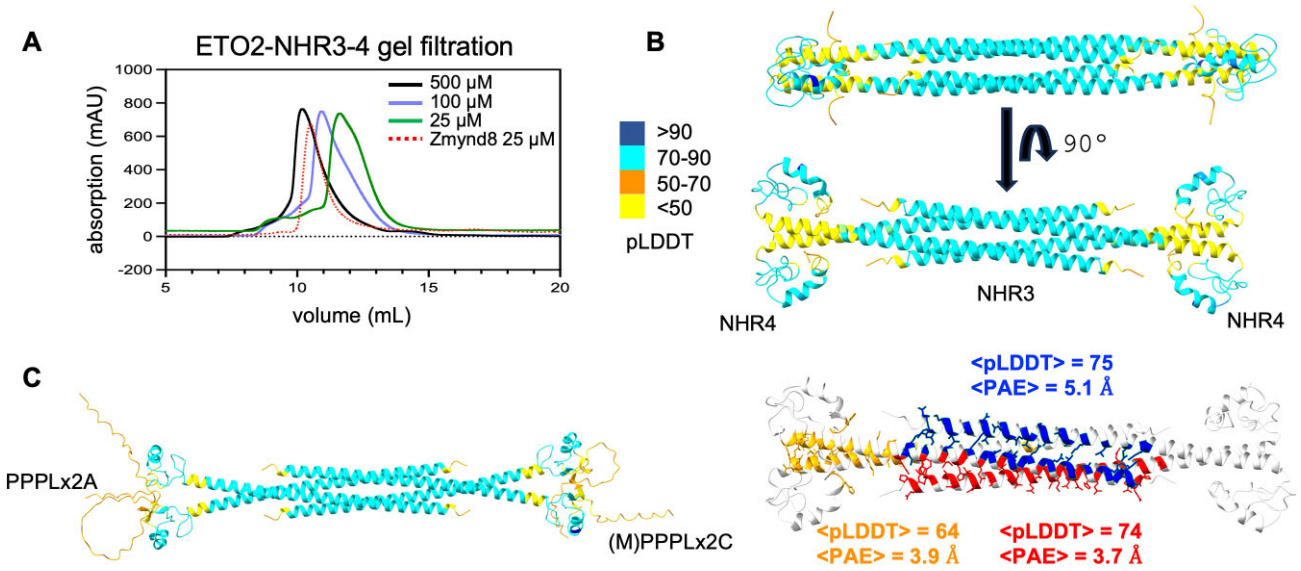


Figure 5. ETO2-NHR3-4 is predicted to form an antiparallel tetramer. (**A**) ETO2-NHR3-4 elutes from a size exclusion column (Superdex75 Increase 10/300, Cytiva) at progressively earlier volumes with increasing concentration of injectant (25, 100, and 500 µM). At the highest concentration, it elutes at a similar position to the Zmynd8 tetramer (25 µM). (**B**) AlphaFold2-multimer models the ETO2-NHR3-4 tetramer as an antiparallel dimer of dimers with predicted local distance difference test (pLDTT) scores in the 70–90 range. The two dimers interact through a well-defined interface (pLDDT = 74 and 75, PAE = 3.7 and 5.1) to form an anti-parallel tetramer (highlighted in red and blue). Extending from the tetrameric interface is a less defined region (pLDTT = 64, PAE = 3.9 A) where the helices interact as parallel dimers (highlighted in orange). (**C**) The AlphaFold2-multimer model of the complex between ETO2-NHR3-4 and PPPLx2A and (**M**)PPPLx2C shows that the bivalent peptides bind to opposite ends of the tetramer, although with much lower confidence scores.

Stabilizing the NHR3 coiled-coil interface promotes the formation of a tetramer

To test the AF2m model of the ETO2-NHR3 domain, we sought to introduce mutations that promote the formation of the antiparallel tetramer. Our goal was to stabilize the complex for functional and biophysical characterization. We subjected the best-ranked AF2m model structures to the point-mutation stabilization protocol (Stabilize Proteins: PM) developed for the Rosetta molecular modeling software [49] as implemented on the ROSIE 2 server [48]. We focused on mutations involving hydrophobic residues at the coiled-coil interface instead of solvent-exposed polar interactions. We modeled potentially stabilizing mutations with AF2m and chose those that generated high-confidence structures for further testing.

To assess the effects of these mutations on tetramer formation, we collected 2D ^{15}N-HSQC spectra of the isolated tetrameric coiled-coil region (amino acids 480–536) for the native and mutant peptides. As shown in Fig. 7A, the spectrum of the native peptide is poorly resolved with broad overlapping peaks, consistent with dynamic exchange between oligomeric states on an intermediate timescale. The Ala501Leu mutation, which improves packing at the coiled-coil interface, significantly improves spectral dispersion and resolution. Introducing a second mutation, Val497Ile, further improves the spectrum. We then measured the circular dichroism spectrum of these peptides to assess helical content. As shown in Fig. 7B, the mutations progressively increase helicity from 29% to 84%, consistent with stabilizing the coiled-coil complex formation. Including these mutations in the AF2m model of the tetramer (Fig. 7C) improves the confidence scores without changing the relative orientation of the domains (compare with Fig. 5B).

We introduced the V497I/A501L mutation into the ETO2-NHR3-4 construct to test the impact of stabilizing the tetrameric coiled-coil on complex formation. We measured binding to PPPLx2A and N-CoRx2A by ITC (Fig. 7C and Supplementary Fig. S8), which shows a similar affinity to the wild-type sequence. Then, we assessed the size of the complex by size exclusion chromatography with multi-angle light scattering (SEC-MALS) (Fig. 7D). The wild-type ETO2-NHR3-4:N-CoRx2A (2:1 molar ratio) complex elutes as a broad peak with variable molecular weight that ranges from that of a dimeric complex (36.5 kDa) to a tetrameric complex (73 kDa) and larger. In contrast, SEC-MALS analysis of the complex with the stabilizing mutations ETO2-NHR3-4 (V497I/A501L):N-CoRx2A (2:1 molar ratio) reveals a single narrow peak with a constant molecular mass of 70 kDa, which is compatible with the tetrameric complex. The results indicate that the wild-type complex is unstable and exists in dynamic exchange between oligomeric states, whereas the double mutation stabilizes the tetrameric complex.

PKA (RIIα) binds to ETO2-NHR3, disrupts oligomerization, and reduces binding to multivalent ligands

In addition to stabilizing the oligomerization of NHR3, we aimed to determine the impact of disrupting the oligomerization domain on the binding of the NHR4 domain to target peptides and γ-globin expression. Previous studies have shown that the cAMP-dependent protein kinase type II-alpha regulatory subunit (PKA (RIIα)) binds directly to the NHR3 domain of ETO with high affinity (K_D = 128 nM) [28]. The structure of this complex shows that PKA (RIIα) binds the predicted tetramerization interface (Fig. 8A and B). Hence, PKA (RIIα) should disrupt ETO2-NHR3 tetramerization and

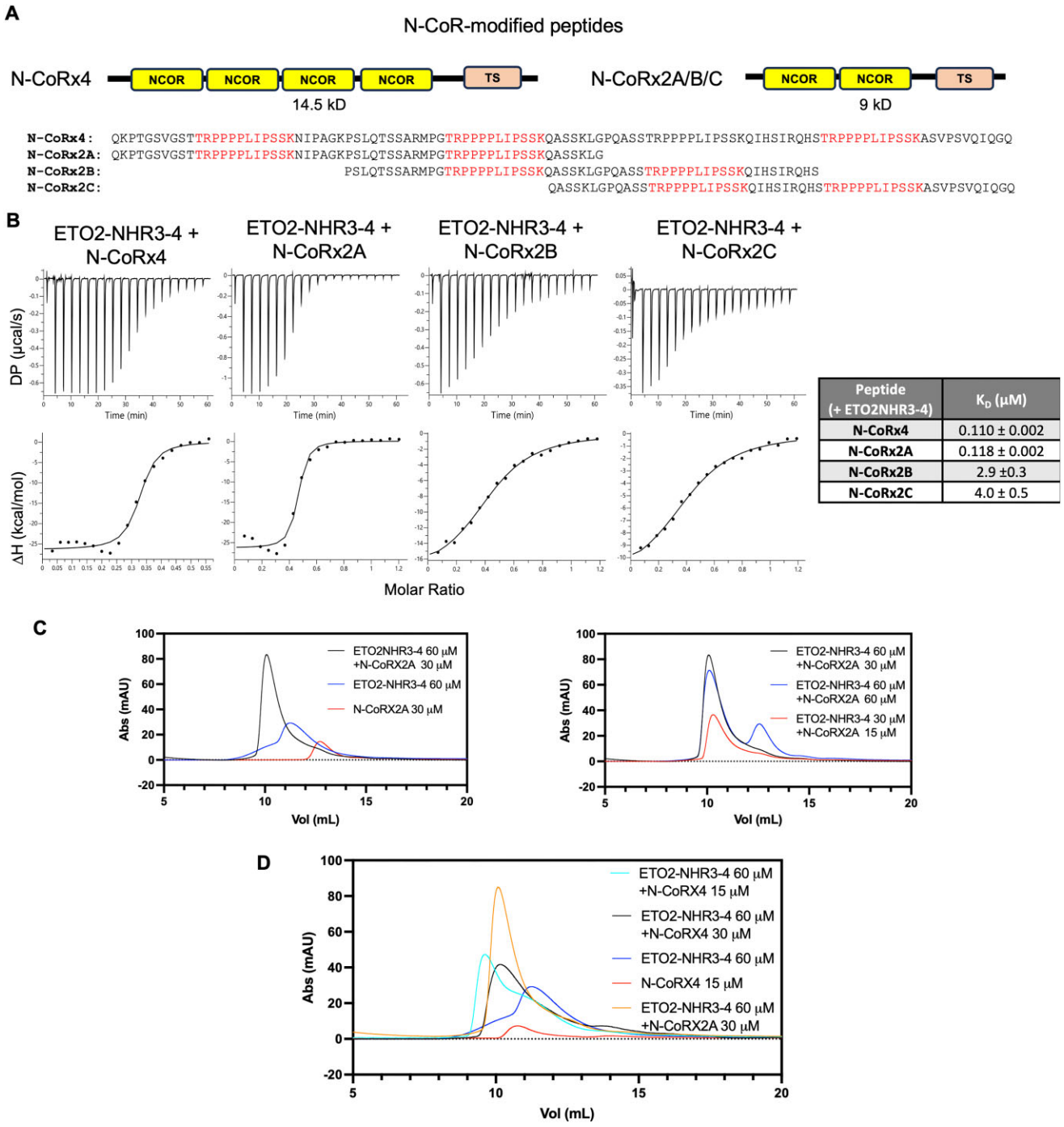
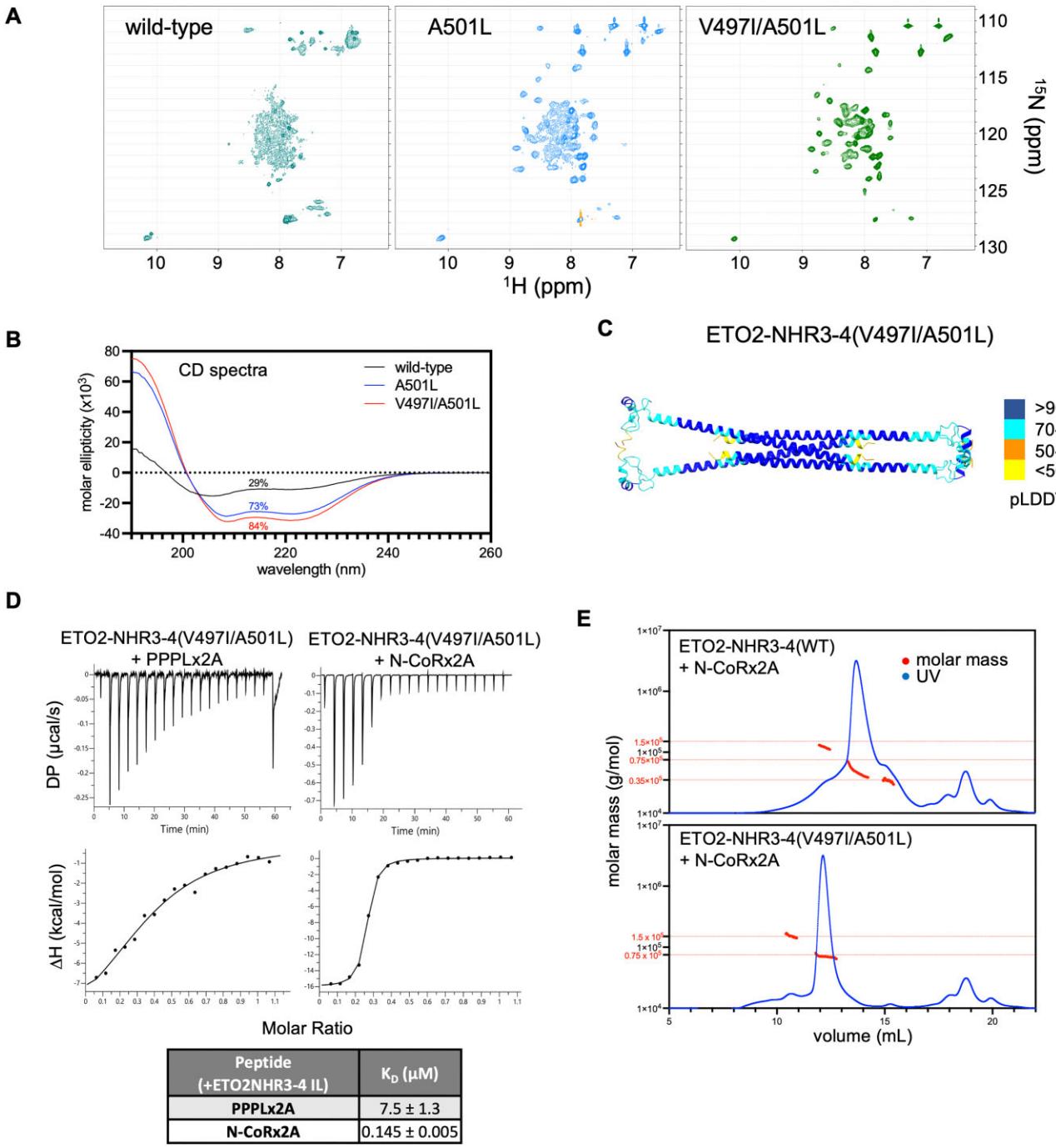
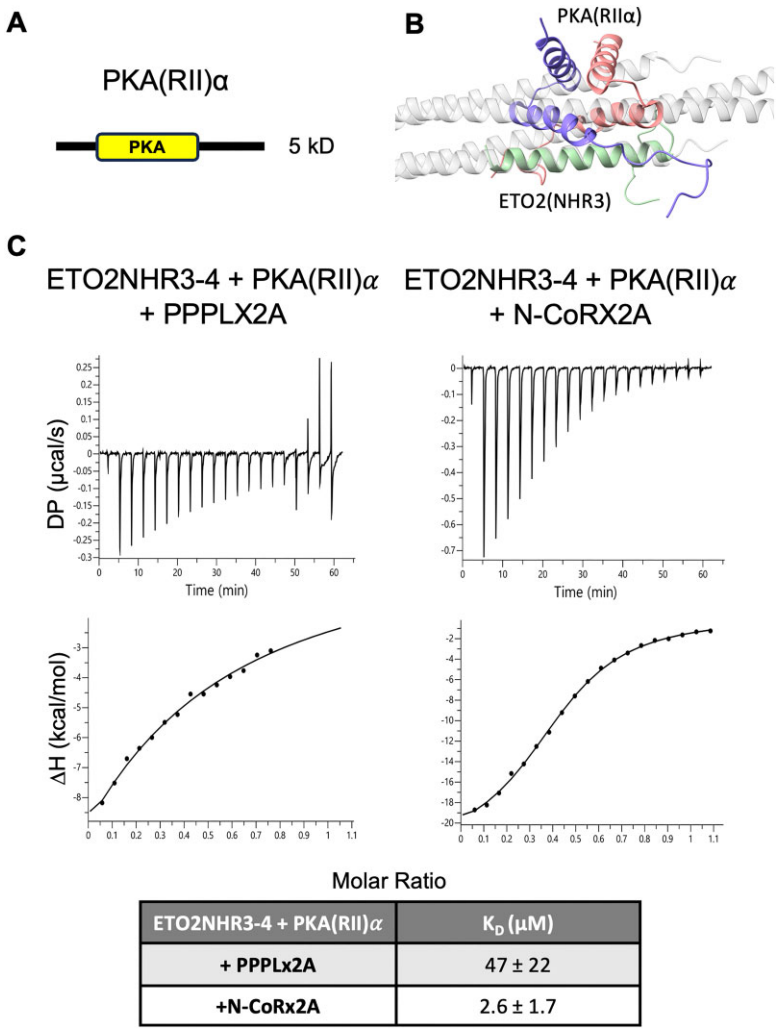


Figure 6. Developing a high-affinity multivalent peptide ligand. (**A**) A diagram depicts the N-CoRx4, N-CoRx2A, N-CoRx2B, and N-CoRx2C peptides and their respective sequences. (**B**) The ITC binding isotherms and associated fit to a one-site binding model are shown for the N-CoRx2A peptides (250 µM) titrated into ETO2-NHR3-4 (50 µM)**.** The reported dissociation constants (K_D) are the average and standard deviation for 2–3 repeated measurements (see Supplementary Fig. S6). (**C**) N-CoRx2A and ETO2-NHR3-4 were combined at 1:2 and 1:1 molar ratios and analyzed by size exclusion chromatography (Superdex-75 Increase 10/300, Cytiva). The mixture of N-CoRx2A and ETO2-NHR3-4 (1:2 molar ratio) elutes earlier than the individual components, consistent with complex formation. With excess N-CoRx2A (1:1 molar ratio), an additional peak elutes at elutes at a position corresponding to unbound N-CoRx2A. (**D**) N-CoRx4 was combined with ETO2-NHR3-4 at a 1:4 and 1:2 ratio and analyzed by size exclusion chromatography (Superdex-75 Increase 10/300, Cytiva). The complex formed between N-CoRx4 and ETO2-NHR3-4 (1:2 molar ratio) elutes at a similar position as N-CoRx2A and ETO2-NHR3-4 (1:2 molar ratio). However, the complex formed between N-CoRx4 and ETO2-NHR3-4 (1:4 molar ratio) elutes at an earlier position (cyan), indicating the formation of a larger oligomeric species.



Peptide (+ETO2NHR3-4 IL)	K_D (µM)
PPPLx2A	7.5 ± 1.3
N-CoRx2A	0.145 ± 0.005

Figure 7. Stabilizing the homo-tetrameric ETO2-NHR3 coiled-coil complex. (**A**) 2D ^{15}N-HSQC spectra of the isolated ^{15}N-ETO2 coiled-coil region (residues 480–536) show that the A501L and V497I/A501L mutations markedly improve peak resolution and spectral dispersion. (**B**) Circular dichroism spectra of the wild-type and mutant ETO2 coiled-coil region show that these mutations increase helical content from 29% (wild-type) to 73% (A501L) and 84% (V497/A501L) (**C**) The ETO2-NHR3-4 (V497I/A501L) mutation markedly improves the overall confidence of the AlphaFold2-multimer model, with pLDDT scores at the tetrameric interface >90 (compare to Fig. 5B). (**D**) The ITC binding isotherms and associated fit to a one-site binding model are shown for the PPPLx2A and N-CoRx2A peptides (250 µM) titrated into ETO2-NHR3-4 (V497I/A501L) (50 µM)**.** The reported dissociation constants (K_D) are the average and standard deviation for two repeated measurements (see Supplementary Fig. S8). (**E**) The complexes formed between N-CoRx2A (60 µM) and ETO2-NHR3-4 (WT) or ETO2-NHR3-4 (V497I/A501L) (120 µM) were analyzed by SEC-MALS. Red horizontal lines are shown at the approximate molecular weights for a dimer (35 kDa), tetramer (75 kDa), and octamer (150 kDa). The sloping molar mass for the wild-type protein indicates a dynamic equilibrium among states. In contrast, the relatively flat molar mass for ETO2-NHR3-4 (V497I/A501L) + N-CoRx2A indicates a more stable complex. Hence, the V497I/A501L mutation stabilizes the formation of the complex, eluting as a single peak with a molecular weight appropriate for a tetramer.



ETO2NHR3-4 + PKA(RII)α	K_D (µM)
+ PPPLx2A	47 ± 22
+N-CoRx2A	2.6 ± 1.7

Figure 8. Disrupting tetramerization by the ETO2-NHR3 domain reduces binding affinity for multivalent ligands. (**A**) A diagram depicts the PKA (RII)α domain used for these studies. (**B**) A ribbon diagram illustrates the structure of the complex between PKA (RII)α dimer (red and blue) and the AML1-ETO-NHR3 domain (green, PDB ID: 2KYG) aligned with the ETO2-NHR3 tetramer model (white), which shows that PKA (RII)α binds to the predicted tetrameric interface. (**C**) The ITC binding isotherms and associated fit to a one-site binding model are shown for the PPPLx2A (see Fig. 4A) and N-CoRx2A (see Fig. 6A) peptides (250 µM) titrated into ETO2-NHR3-4 (50 µM) in the presence of PKA (RII)α (100 µM). The reported dissociation constants (K_D) are the average and standard deviation for two repeated measurements (see Supplementary Fig. S9).

reduce binding to multivalent ligands. To test this hypothesis, we measured the binding affinity between ETO2-NHR3-4 and N-CoRx2A in the presence of PKA (RIIα) (Fig. 8C and Supplementary Fig. S9). The results show that PKA (RIIα) disrupts the binding of ETO2-NHR3-4 to N-CoRX2A, reducing the affinity approximately 30-fold ($K_D = 3 \pm 2$ µM). Likewise, it reduced the binding affinity for the PPPLx2A peptide approximately 10-fold ($K_D = 50 \pm 20$ µM).

Targeted disruption of ETO2-NHR3-4 increases γ-globin gene expression and induces differentiation of HUDEP-2 adult phenotype human erythroid cells as well as K562 leukemia cells. Having established a method to disrupt ETO2 interactions with multivalent peptides, we sought to test the impact of this disruption on globin gene regulation. As we have previously demonstrated the critical role of NuRD in silencing the γ-globin in adult erythroid cells [18, 22, 42, 59–61], we tested the enforced expression of N-CoRx2A in the human-derived erythroid progenitor cell line 2 (HUDEP-2) that faithfully recapitulates human hemoglobin regulation. For these experiments, an N-CoRx2A construct containing a nuclear localization signal was cloned into a lentiviral expression vector under the control of the strong SFFV promoter. The results show that enforced expression of this peptide induces fetal γ−globin gene expression (Fig. 9). In addition, we tested the enforced expression of N-CoRX2A in K562 cells to directly compare our results with the previous studies, which showed that depletion of ETO2 in K562 leukemia cells induces γ-globin expression. [3] As shown in Supplementary Fig. S10, the enforced expression of N-CoRx2A in K562 cells induced the expression of γ-globin RNA to a level ~ 50% as great as was previously reported after complete knockout of ETO2 in this cell line [3].

Recent work showed that the knockout of ETO2 protein in leukemia cell lines induces differentiation, which has proven to be a potent therapeutic strategy for treating acute myeloid leukemia, as demonstrated by the effectiveness of all-trans retinoic acid for the treatment of acute promyelocytic leukemia [15] and the more recently developed menin inhibitors [62]. Therefore, we tested whether disrupting ETO2-NHR3-4 would induce differentiation in HUDEP-2 erythroid

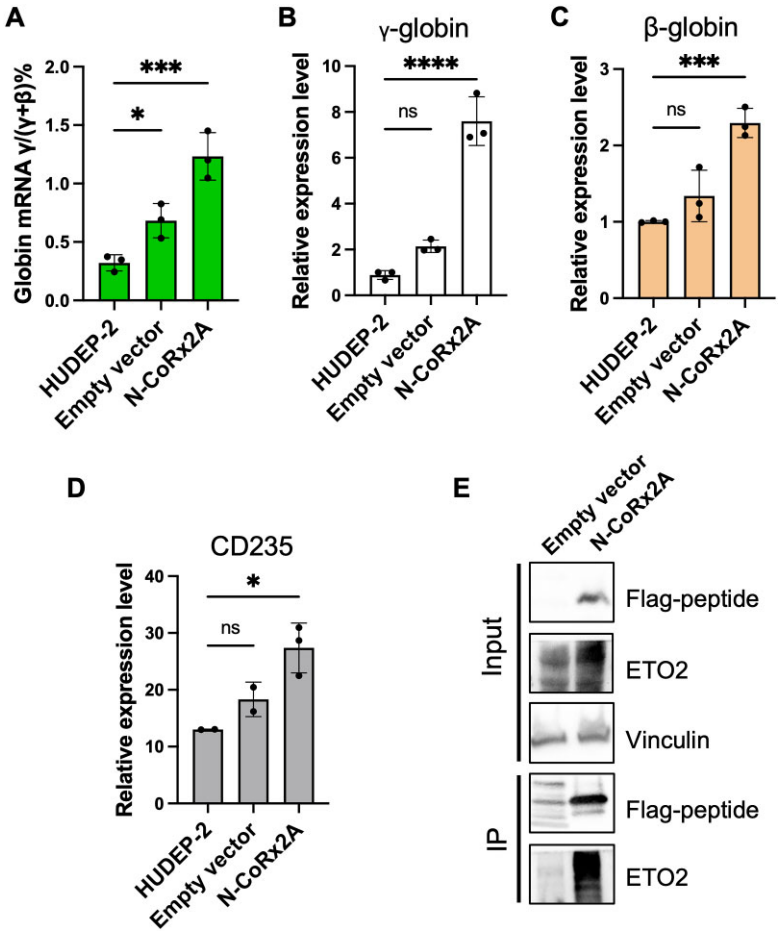


Figure 9. Enforced expression of N-CoRx2A induces g-globin gene expression and differentiation in HUDEP-2 cells. (**A**) Q-PCR results show the γ/ (γ + β) ratio of mRNA in HUDEP-2 cells expressing N-CoRx2A peptide. (**B**) and (**C**) Q-PCR results of γ-globin and β-globin mRNA, respectively. (**D**) Q-PCR results of CD235 mRNA in N-CoRx2A peptide expressing HUDEP-2 cells. The results are shown as the mean ± SD, $n = 3$. *P* value was calculated by one-way ANOVA. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. (**E**) Co-immunoprecipitation shows the interaction of expressed Flag-tagged N-CoRx2A with endogenous ETO2 in HUDEP-2 cells.

progenitor cells and K562 leukemia cells. By measuring CD235 and β-globin gene expression in HUDEP-2 cells (Fig. 9), we found that enforced expression of N-CoRx2A resulted in erythroid differentiation. Similarly, N-CoRx2A induced differentiation of K562 cells as measured by qPCR assay of expression of CD235 (Supplementary Fig. S10). Further, immunoprecipitation of Flag-tagged N-CoRx2A in nuclear extracts followed by western blotting confirmed that the peptide binds strongly to both ETO2 and ETO in both cell lines (Fig. 9 and Supplementary Fig. S10). These studies show that targeted disruption ETO protein complexes with a high-affinity multivalent ligand augments fetal hemoglobin expression and induces differentiation. Hence, further development of more potent multivalent inhibitors could provide novel molecular probes that selectively disrupt the biological function of ETO proteins, which may be beneficial therapeutically.

Discussion

The ETO proteins function as structural components that bridge between transcription factors and co-regulatory complexes, typically silencing transcription of the associated gene. In hematopoietic cells, the ETO proteins are found in a large complex that includes the T-cell acute lymphoblastic leukemia transcription factor (TAL1), GATA binding protein 1 (GATA1), LIM-only protein 2 (LMO2), and LIM domain binding protein 1 (LDB1) [1, 4, 16, 63]. This complex has multiple dimerization domains that contribute to the formation of chromatin loops between enhancers and promoters. The relative stoichiometry of ETO proteins at a particular locus correlates with gene silencing [57]. Hence, the ability to oligomerize appears to play a critical role in their function within the larger complex.

This work investigates how NHR3, a relatively understudied coiled-coil domain in ETO2, might contribute to the multivalent recruitment of the NuRD complex. This domain immediately precedes the NHR4, which directly binds to co-regulatory proteins. Our results support a model in which the NHR3 region forms an antiparallel tetramer that increases binding affinity for bivalent ligands with the appropriate spacing and linker composition between binding sites. Notably, though, the NHR3 only very weakly oligomerizes in isolation.

In the context of the whole protein, the stable antiparallel tetrameric NHR2 domain [26, 27] promotes tetramerization of the NHR3 domain, such that both domains contribute to the high-affinity binding of bivalent ligands. This cooperativity reflects an effective increase in the local concentration

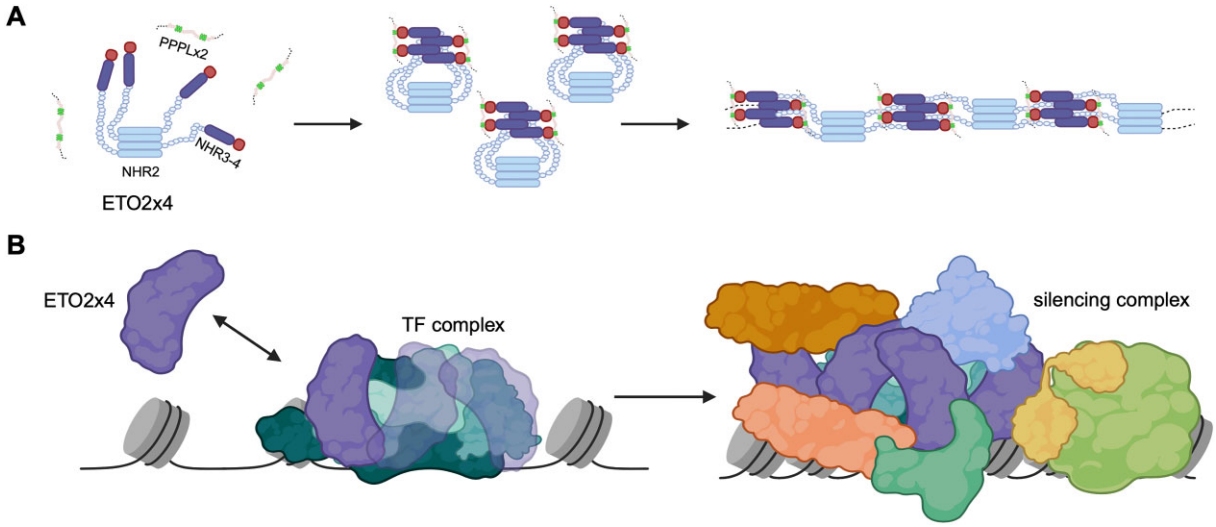


Figure 10. Model of ETO2 oligomerization in transcription regulation. (**A**) ETO2 forms a stable tetramer through its NHR2 domain (light blue). Binding to a multivalent ligand (PPPLx2) stabilizes antiparallel tetramer formation of the NHR3-4 domains (purple). At high density, the combination of stable (NHR2) and dynamic (NHR3) tetrameric domains leads to the formation of large arrays of the ETO2 tetramer. (**B**) When ETO2 (purple) is recruited by transcription factor complexes to a genetic locus at sufficient density, it can start forming these arrays that recruit more co-regulatory protein complexes (e.g. NuRD, N-CoR, and SMRT), ultimately leading to a large silencing complex and compacted chromatin.

of NHR3 domains through tethering to the stably tetrameric NHR2 domain. We estimate the average end-to-end distance of the 38 amino acid linker separating the two domains as approximately 62 Å, calculated from $(C_n nl^2)^{1/2}$, where C_n is the characteristic ratio (~7.5 for 36 amino acids), n is the number of amino acids, and l is 3.8 Å [64]. Therefore, the effective local concentration (one NHR3 domain in a sphere of radius 62 Å) is approximately 1.6 mM, well above that necessary to drive NHR3 tetramerization. Consequently, the weak NHR3 self-association creates a dynamic interaction that can be modulated as a regulatory mechanism to control ETO2 function.

Structural model of the ETO2-NHR3 domain self-association

The very weak self-association of NHR3 makes traditional structural studies very challenging. Hence, to better understand the function of NHR3, we used a combination of biophysical studies, NMR structural analysis, structure prediction with AF2m, molecular modeling with Rosetta, and in-cell analyses. This combined approach allowed us to generate and test a structural model of the ETO2-NHR3-4 domains. The results show that the domain likely forms an antiparallel tetramer such that a pair of NHR4 domains at either end interact with bivalent ligands. This arrangement explains why ETO2-NHR3-4 binds with equivalent affinity to a peptide containing two or four binding sites. In addition, this arrangement indicates that the spacing and linker connecting binding motifs contribute to target selectivity.

Polyproline binding domains characterize large families of protein domains (SH3, GYF, EVH1, and WW) involved in gene regulation and signaling [65–68]. These domains typically recognize the polyproline type II helical structure adopted by the individual polyproline motifs with relatively weak affinity. Hence, the potential binding sites share significant overlap between the domain families. This degeneracy raises a question of how the domains target specific proteins for function. Multivalency helps to both improve target binding affinity and provide a mechanism for increased selectivity. As we found here, the spacing and linker sequence can significantly impact the binding affinity. ETO2-NHR3-4 preferentially binds to the first two binding sites in GATAD2A even when the individual motifs are identical. This observation suggests that a longer linker between sites is necessary, correlating with the physical separation of NHR4 domains in the AF2m model.

Having established that ETO2-NHR3 tetramerization stabilizes binding to multivalent ligands, we used this information to disrupt ETO2 function in cells. We developed the N-CoRx2A peptide that binds to ETO2 much more tightly than the native GATAD2A or any other polyproline binding site reported. As predicted, this peptide disrupted fetal hemoglobin gene silencing and induced differentiation of leukemia cells. These studies confirm the functional importance of multivalent interactions for the ETO proteins and open the door to the future development of molecules that inhibit normal and aberrant gene silencing for therapeutic benefit.

Model for ETO2 recruitment of coregulatory proteins

Based on this work, we propose a model in which co-regulatory complex recruitment by ETO proteins depends on multivalent association and NHR3-tetramerization. Factors that modulate this self-association influence which co-factors are stably recruited, which then dictates activation or silencing of the transcription of the associated gene. This model helps explain prior domain mapping experiments, which often implicate the NHR2 domain as directly binding to co-regulatory complexes. In our model, deleting NHR2 indirectly disrupts binding by preventing stable tetramerization. Selectively deleting NHR4 does not prevent stable tetramerization with native proteins through NHR2. Therefore, the heteromeric complexes comprising native and NHR4-deleted proteins can interact with co-regulatory complexes through the native NHR4. The NHR3 domain, in isolation, however,

does not form a sufficiently stable tetramer to drive multivalent association with target proteins.

Furthermore, this model helps explain why the apparent stoichiometry of ETO proteins at a particular locus correlates with gene activation or silencing. For example, the relative ratio of NLI/ETO2 complexes at the γ-globin locus, not simply the presence and absence of ETO2, correlates with fetal hemoglobin expression in K562 cells [16]. Effective recruitment of silencing co-regulatory complexes, such as NuRD, requires multiple copies of ETO proteins to interact with multivalent binding motifs. Therefore, gene silencing correlates with numerous copies of ETO proteins at a given locus. As we found, the (M)PPPLx4 binding region in GATAD2A can crosslink multiple ETO2 tetramers, ultimately forming very large and dynamic complexes that compact and silence chromatin (Fig. 10).

In summary, we have shown that ETO2 binds to the GATAD2A component of NuRD through multivalent interaction between the NHR4 domain and polyproline-leucine motifs in GATAD2A. Multivalent binding is critical for gene silencing and depends on the weak self-association of the ETO2-NHR3 coiled-coil domain. Together, these interactions provide a mechanism for selective target engagement and dynamic regulation of gene silencing and hematopoietic cell differentiation, and they offer the opportunity to develop novel multivalent inhibitory molecules that selectively disrupt ETO2 function.

Acknowledgements

We thank the support of the University of North Carolina Macromolecular Interactions Facility and NMR Facility. Graphical abstract and Fig. 10 are created with Biorender.com.

Author contributions; D.W., G.G., and G.D. conceived the original ideas, designed the project, and wrote the manuscript with inputs from M.W., G.L., C.T., and S.S. G.D., D.W., G.L, T.S., and S.S. performed the majority of the experiments with participation from P.A, O.A, and T.L.

Supplementary data

Supplementary data are available at NAR online.

Conflict of interest

None declared.

Funding

National Institute of Diabetes, Digestive, and Kidney Diseases [R01DK115563 to D.W., G.G, R56DK29902 to G.G.], National Cancer Institute [P30CA16086 to UNC Lineberger Cancer Center, P30CA016059 to VCU Massey Cancer Center], National Institute of General Medical Sciences [R01GM118499 and R35GM145227 to M.W.]. The authors acknowledge the generous support provided by Eshelman Innovation at the UNC Eshelman School of Pharmacy and the UNC Lineberger Comprehensive Cancer Center (LCCC). Funding for open access charge: National Institute of Diabetes, Digestive, and Kidney Diseases (to D.W. and G.G.).

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and/or its supplementary data or can be made available upon reasonable request. The NMR assignments and solution structure of the ETO2-NHR4:MPPL-3 complex have been deposited in the BMRB (accession code: 31200) and RCSB (PDBID: 9DE2), respectively.

References

1. Fujiwara T, Lee H-Y, Sanalkumar R *et al.* Building multifunctionality into a complex containing master regulators of hematopoiesis. *Proc Natl Acad Sci USA* 2010;**107**:20429–34. https://doi.org/10.1073/pnas.1007804107
2. Goardon N, Lambert JA, Rodriguez P *et al.* ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. *EMBO J* 2006;**25**:357–66. https://doi.org/10.1038/sj.emboj.7600934
3. Guo X, Plank-Bazinet J, Krivega I *et al.* Embryonic erythropoiesis and hemoglobin switching require transcriptional repressor ETO2 to modulate chromatin organization. *Nucleic Acids Res* 2020;**48**:10226–40. https://doi.org/10.1093/nar/gkaa736
4. Hoang T, Lambert JA, Martin R. SCL/TAL1 in hematopoiesis and cellular reprogramming. *Curr Top Dev Biol* 2016;**118**:163–204.
5. Davis JN, McGhee L, Meyers S. The ETO (MTG8) gene family. *Gene* 2003;**303**:1–10. https://doi.org/10.1016/S0378-1119(02)01172-1
6. Downing JR. The AML1-ETO chimaeric transcription factor in acute myeloid leukaemia: biology and clinical significance. *Br J Haematol* 1999;**106**:296–308. https://doi.org/10.1046/j.1365-2141.1999.01377.x
7. Lin S, Mulloy JC, Goyama S. RUNX1-ETO leukemia. *Adv Exp Med Biol* 2017;**962**:151–73.
8. Gelmetti V, Zhang J, Fanelli M *et al.* Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. *Mol Cell Biol* 1998;**18**:7185–91. https://doi.org/10.1128/MCB.18.12.7185
9. Okuda T, Cai Z, Yang S *et al.* Expression of a knocked-in AML1-ETO leukemia gene inhibits the establishment of normal definitive hematopoiesis and directly generates dysplastic hematopoietic progenitors. *Blood* 1998;**91**:3134–43. https://doi.org/10.1182/blood.V91.9.3134
10. Miyoshi H, Kozu T, Shimizu K *et al.* The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. *EMBO J* 1993;**12**:2715–21. https://doi.org/10.1002/j.1460-2075.1993.tb05933.x
11. Erickson P, Gao J, Chang KS *et al.* Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, AML1/ETO, with similarity to Drosophila segmentation gene, runt. *Blood* 1992;**80**:1825–31. https://doi.org/10.1182/blood.V80.7.1825.1825
12. Thirant C, Lopez C, Malinge S *et al.* Molecular pathways driven by ETO2-GLIS2 in aggressive pediatric leukemia. *Mol Cell Oncol* 2017;**4**:e1345351. https://doi.org/10.1080/23723556.2017.1345351
13. Piqué-Borràs M-R, Jevtic Z, Bagger FO *et al.* The NFIA-ETO2 fusion blocks erythroid maturation and induces pure erythroid leukemia in cooperation with mutant TP53. *Blood* 2023;**141**:2245–60.
14. Micci F, Thorsen J, Panagopoulos I *et al.* High-throughput sequencing identifies an NFIA/CBFA2T3 fusion gene in acute erythroid leukemia with t(1;16)(p31;q24). *Leukemia* 2013;**27**:980–2. https://doi.org/10.1038/leu.2012.266
15. Steinauer N, Guo C, Zhang J. The transcriptional corepressor CBFA2T3 inhibits all-trans-retinoic acid-induced myeloid gene expression and differentiation in acute myeloid leukemia. *J Biol*

Chem 2020;**295**:8887–900. https://doi.org/10.1074/jbc.RA120.013042
16. Kiefer CM, Lee J, Hou C *et al.* Distinct Ldb1/NLI complexes orchestrate gamma-globin repression and reactivation through ETO2 in human adult erythroid cells. *Blood* 2011;**118**:6200–8. https://doi.org/10.1182/blood-2011-06-363101
17. Lenz J, Brehm A. Conserved mechanisms of NuRD function in hematopoetic gene expression. *Enzymes* 2023;**53**:7–32.
18. Shang S, Li X, Azzo A *et al.* MBD2a–NuRD binds to the methylated γ-globin gene promoter and uniquely forms a complex required for silencing of HbF expression. *Proc Natl Acad Sci USA* 2023;**120**:e2302254120. https://doi.org/10.1073/pnas.2302254120
19. Liang Y, Zhang X, Liu Y *et al.* GATA zinc finger domain-containing protein 2A (GATAD2A) deficiency reactivates fetal haemoglobin in patients with beta-thalassaemia through impaired formation of methyl-binding domain protein 2 (MBD2)-containing nucleosome remodelling and deacetylation (NuRD) complex. *Br J Haematol* 2021;**193**:1220–27. https://doi.org/10.1111/bjh.17511
20. Kransdorf EP, Wang SZ, Zhu SZ *et al.* MBD2 is a critical component of a methyl cytosine-binding protein complex isolated from primary erythroid cells. *Blood* 2006;**108**:2836–45. https://doi.org/10.1182/blood-2006-04-016394
21. Rupon JW, Wang SZ, Gnanapragasam M *et al.* MBD2 contributes to developmental silencing of the human epsilon-globin gene. *Blood Cells Mol Dis* 2011;**46**:212–9. https://doi.org/10.1016/j.bcmd.2011.01.001
22. Rupon JW, Wang SZ, Gaensler K *et al.* Methyl binding domain protein 2 mediates gamma-globin gene silencing in adult human betaYAC transgenic mice. *Proc Natl Acad Sci USA* 2006;**103**:6617–22. https://doi.org/10.1073/pnas.0509322103
23. Masuda T, Wang X, Maeda M *et al.* Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science* 2016;**351**:285–9. https://doi.org/10.1126/science.aad3312
24. Sher F, Hossain M, Seruggia D *et al.* Rational targeting of a NuRD subcomplex guided by comprehensive in situ mutagenesis. *Nat Genet* 2019;**51**:1149–59. https://doi.org/10.1038/s41588-019-0453-4
25. Park S, Chen W, Cierpicki T *et al.* Structure of the AML1-ETO eTAFH domain-HEB peptide complex and its contribution to AML1-ETO activity. *Blood* 2009;**113**:3558–67. https://doi.org/10.1182/blood-2008-06-161307
26. Liu Y, Cheney MD, Gaudet JJ *et al.* The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. *Cancer Cell* 2006;**9**:249–60. https://doi.org/10.1016/j.ccr.2006.03.012
27. Sun XJ, Wang Z, Wang L *et al.* A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* 2013;**500**:93–7. https://doi.org/10.1038/nature12287
28. Corpora T, Roudaia L, Oo ZM *et al.* Structure of the AML1-ETO NHR3-PKA(RIIalpha) complex and its contribution to AML1-ETO activity. *J Mol Biol* 2010;**402**:560–77. https://doi.org/10.1016/j.jmb.2010.08.007
29. Donghui WU, Haitao Y, Xiaoyu XUE *et al.* Oligomerization study of NHR3 and NHR4 domains from ETO protein involved in t(8;21)-associated acute myeloid leukemia. *Chin Sci Bull* 2005;**50**:875–79. https://doi.org/10.1007/BF02897381
30. Wang J, Hoshino T, Redner RL *et al.* ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. *Proc Natl Acad Sci USA* 1998;**95**:10860–5. https://doi.org/10.1073/pnas.95.18.10860
31. Lutterbach B, Westendorf JJ, Linggi B *et al.* ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Mol Cell Biol* 1998;**18**:7176–84. https://doi.org/10.1128/MCB.18.12.7176
32. Liu Y, Chen W, Gaudet J *et al.* Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. *Cancer Cell* 2007;**11**:483–97. https://doi.org/10.1016/j.ccr.2007.04.010
33. Harter MR, Liu CD, Shen CL *et al.* BS69/ZMYND11 C-terminal domains bind and inhibit EBNA2. *PLoS Pathog* 2016;**12**:e1005414. https://doi.org/10.1371/journal.ppat.1005414
34. Spruijt CG, Luijsterburg MS, Menafra R *et al.* ZMYND8 Co-localizes with NuRD on target genes and regulates poly(ADP-Ribose)-dependent recruitment of GATAD2A/NuRD to sites of DNA damage. *Cell Rep* 2016;**17**:783–98. https://doi.org/10.1016/j.celrep.2016.09.037
35. Gong F, Clouaire T, Aguirrebengoa M *et al.* Histone demethylase KDM5A regulates the ZMYND8-NuRD chromatin remodeler to promote DNA repair. *J Cell Biol* 2017;**216**:1959–74. https://doi.org/10.1083/jcb.201611135
36. Chen Y, Tong X, Lu R *et al.* All-trans retinoic acid in hematologic disorders: not just acute promyelocytic leukemia. *Front Pharmacol* 2024;**15**:1404092. https://doi.org/10.3389/fphar.2024.1404092
37. Schmidt TGM, Batz L, Bonet L *et al.* Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr Purif* 2013;**92**:54–61. https://doi.org/10.1016/j.pep.2013.08.021
38. Delaglio F, Grzesiek S, Vuister GW *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 1995;**6**:277–93. https://doi.org/10.1007/BF00197809
39. Skinner SP, Fogh RH, Boucher W *et al.* CcpNmr AnalysisAssign: A flexible platform for integrated NMR analysis. *J Biomol NMR* 2016;**66**:111–24. https://doi.org/10.1007/s10858-016-0060-y
40. Rückert M, Otting G. Alignment of biological macromolecules in novel nonionic liquid crystalline Media for NMR experiments. *J Am Chem Soc* 2000;**122**:7793–7. https://doi.org/10.1021/ja001068h
41. Walavalkar NM, Gordon N, Williams DC Jr. Unique features of the anti-parallel, heterodimeric coiled-coil interaction between methyl-cytosine binding domain 2 (MBD2) homologues and GATA zinc finger domain containing 2A (GATAD2A/p66alpha). *J Biol Chem* 2013;**288**:3419–27. https://doi.org/10.1074/jbc.M112.431346
42. Gnanapragasam MN, Scarsdale JN, Amaya ML *et al.* p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proc Natl Acad Sci USA* 2011;**108**:7487–92. https://doi.org/10.1073/pnas.1015341108
43. Desai M, Webb HD, Sinanan LM *et al.* An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex. *Nucleic Acids Res* 2015;**43**:3100–13. https://doi.org/10.1093/nar/gkv168
44. Evans R, O'Neill M, Pritzel A *et al.* Protein complex prediction with AlphaFold-multimer. bioRxiv, https://doi.org/10.1101/2021.10.04.463034, 10 March 2022, preprint: not peer reviewed.
45. Jumper J, Evans R, Pritzel A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583–9. https://doi.org/10.1038/s41586-021-03819-2
46. Meng EC, Goddard TD, Pettersen EF *et al.* UCSF ChimeraX: tools for structure building and analysis. *Protein Sci* 2023;**32**:e4792. https://doi.org/10.1002/pro.4792
47. Genz LR, Mulvaney T, Nair S *et al.* PICKLUSTER: a protein-interface clustering and analysis plug-in for UCSF ChimeraX. *Bioinformatics* 2023;**39**:badt629. https://doi.org/10.1093/bioinformatics/btad629
48. Lyskov S, Chou F-C, Conchúir SÓ *et al.* Serverification of molecular modeling applications: the Rosetta Online server that includes everyone (ROSIE). *PLoS One* 2013;**8**:e63906. https://doi.org/10.1371/journal.pone.0063906
49. Thieker DF, Maguire JB, Kudlacek ST *et al.* Stabilizing proteins, simplified: a Rosetta-based webtool for predicting favorable

mutations. *Protein Sci* 2022;**31**:e4428. https://doi.org/10.1002/pro.4428
50. Jackson DY, King DS, Chmielewski J *et al.* General-approach to the synthesis of short alpha-helical peptides. *J Am Chem Soc* 1991;**113**:9391–2. https://doi.org/10.1021/ja00024a067
51. Andrews NC, Faller DV. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res* 1991;**19**:2499. https://doi.org/10.1093/nar/19.9.2499
52. Chen-Wichmann L, Shvartsman M, Preiss C *et al.* Compatibility of RUNX1/ETO fusion protein modules driving CD34+ human progenitor cell expansion. *Oncogene* 2019;**38**:261–72. https://doi.org/10.1038/s41388-018-0441-7
53. Spadaccini R, Perrin H, Bottomley MJ *et al.* Structure and functional analysis of the MYND domain. *J Mol Biol* 2006;**358**:498–508. https://doi.org/10.1016/j.jmb.2006.01.087
54. Sirinupong N, Brunzelle J, Ye J *et al.* Crystal structure of cardiac-specific histone methyltransferase SmyD1 reveals unusual active site architecture. *J Biol Chem* 2010;**285**:40635–44. https://doi.org/10.1074/jbc.M110.168187
55. Sirinupong N, Brunzelle J, Doko E *et al.* Structural insights into the autoinhibition and posttranslational activation of histone methyltransferase SmyD3. *J Mol Biol* 2011;**406**:149–59. https://doi.org/10.1016/j.jmb.2010.12.014
56. Jiang Y, Sirinupong N, Brunzelle J *et al.* Crystal structures of histone and p53 methyltransferase SmyD2 reveal a conformational flexibility of the autoinhibitory C-terminal domain. *PLoS One* 2011;**6**:e21640. https://doi.org/10.1371/journal.pone.0021640
57. Tatko CD, Waters ML. Investigation of the nature of the methionine-pi interaction in beta-hairpin peptide model systems. *Protein Sci* 2004;**13**:2515–22. https://doi.org/10.1110/ps.04820104
58. Wichmann C, Becker Y, Chen-Wichmann L *et al.* Dimer-tetramer transition controls RUNX1/ETO leukemogenic activity. *Blood* 2010;**116**:603–13. https://doi.org/10.1182/blood-2009-10-248047
59. Amaya M, Desai M, Gnanapragasam MN *et al.* Mi2beta-mediated silencing of the fetal gamma-globin gene in adult erythroid cells. *Blood* 2013;**121**:3493–501. https://doi.org/10.1182/blood-2012-11-466227
60. Yu X, Azzo A, Bilinovich SM *et al.* Disruption of the MBD2-NuRD complex but not MBD3-NuRD induces high level HbF expression in human adult erythroid cells. *Haematologica* 2019;**104**:2361–71. https://doi.org/10.3324/haematol.2018.210963
61. Leighton GO, Shang S, Hageman S *et al.* Analysis of the complex between MBD2 and the histone deacetylase core of NuRD reveals key interactions critical for gene silencing. *Proc Natl Acad Sci USA* 2023;**120**:e2307287120. https://doi.org/10.1073/pnas.2307287120
62. Fiskus W, Mill CP, Birdwell C *et al.* Targeting of epigenetic co-dependencies enhances anti-AML efficacy of Menin inhibitor in AML with MLL1-r or mutant NPM1. *Blood Cancer J* 2023;**13**:53. https://doi.org/10.1038/s41408-023-00826-6
63. Chen Q, Cevher MA, Jiang Q *et al.* LYL1 facilitates AETFC assembly and gene activation by recruiting CARM1 in t(8;21) AML. *Proc Natl Acad Sci USA* 2022;**119**:e2213718119. https://doi.org/10.1073/pnas.2213718119
64. Cantor CR, Schimmel PR. Configurational statistics of polymer chains. In: *Biophysical Chemistry: Part III: The Behavior of Biological Macromolecules*. W. H. Freeman and Company, New York, New York, USA, 1980, pp. 979–1018.
65. Kurochkina N, Guha U. SH3 domains: modules of protein-protein interactions. *Biophys Rev* 2013;**5**:29–39. https://doi.org/10.1007/s12551-012-0081-z
66. Martinez JC, Castillo F, Ruiz-Sanz J *et al.* Understanding binding affinity and specificity of modular protein domains: a focus in ligand design for the polyproline-binding families. *Adv Protein Chem Struct Biol* 2022;**130**:161–88.
67. Fedorov AA, Fedorov E, Gertler F *et al.* Structure of EVH1, a novel proline-rich ligand-binding module involved in cytoskeletal dynamics and neural function. *Nat Struct Biol* 1999;**6**:661–6. https://doi.org/10.1038/10717
68. Hwang T, Parker SS, Hill SM *et al.* Native proline-rich motifs exploit sequence context to target actin-remodeling Ena/VASP protein ENAH. *eLife* 2022;**11**:e70680. https://doi.org/10.7554/eLife.70680

Received: September 9, 2024. **Revised:** April 4, 2025. **Editorial Decision:** May 6, 2025. **Accepted:** May 12, 2025