Communication

cMYC protein interactions and chromatin association in NUT carcinoma

Barry M. Zee[1,2#], Artyom A. Alekseyenko[1,2#], Anne E. Smolko[1,2], Hyuckjoon Kang[1,2], Mitzi I. Kuroda[1,2]*

*Corresponding author

#Equal contribution

[1]Div. of Genetics, Dept. of Medicine, Brigham and Women's Hospital, Boston, MA.

[2]Dept. of Genetics, Harvard Medical School, Boston, MA.

Simple Summary:

A longstanding goal in biology is to understand how protein interactions influence and reflect cellular disease states. MYC is a critical regulator of cellular proliferation that is mis-regulated in many cancers, including those with NUTM1-rearrangements featured in this issue. NUT carcinoma cells are dependent on MYC expression, which blocks differentiation. Using a crosslinking approach to identify MYC protein interactions in a NUTM1-rearranged patient cell line, we found that MYC interacts primarily with subunits of the NuA4 lysine acetyltransferase (KAT5), one of two KAT complexes previously discovered as MYC interactors in non-NUTM1 cell lines.

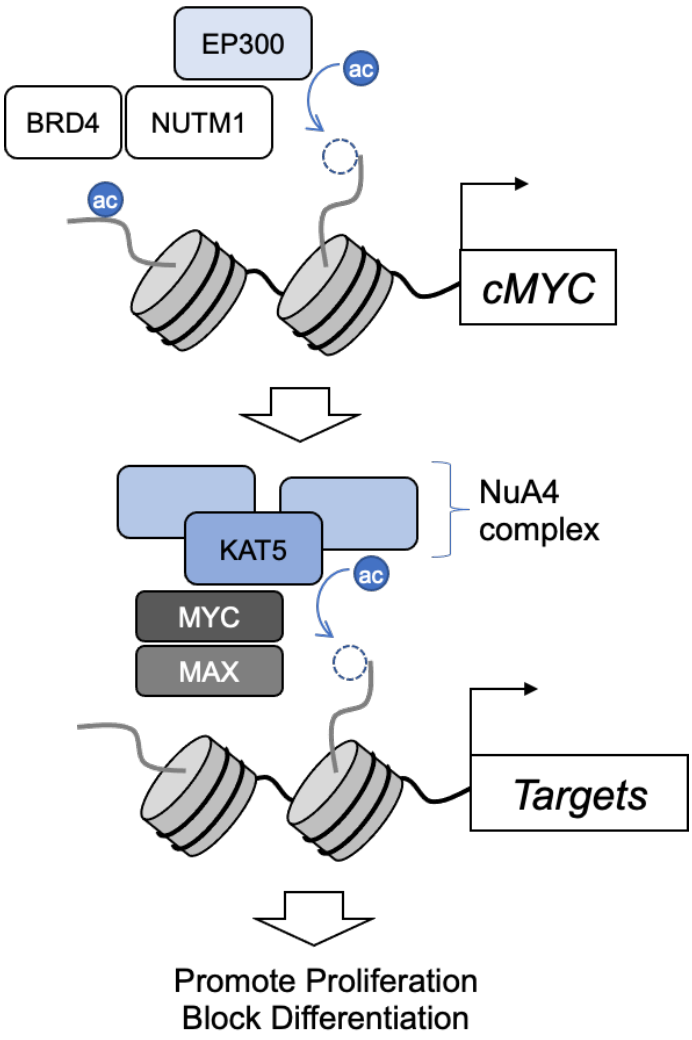
Abstract:

The oncogene c*MYC* (HGNC:7553) is a critical genomic target of the BRD4-NUT (B4N) protein that defines many of the NUTM1-rearrangement cancer subtypes in NUT carcinoma (NC). Previously, we reported that B4N interacts with the EP300 lysine acetyltransferase (KAT3B) and creates hyperacetylated "megadomains" that activate downstream genes such as *cMYC*. Here we ask how misregulated cMYC in turn interacts with protein partners and target genes in patient-derived NC797 cells, and whether these interactions change in response to B4N inactivation. We used CRISPR-Cas9 mediated knock-in of a BioTAP affinity tag to analyze cMYC protein expressed from its normal chromosomal context. This allowed us to implement a crosslinking purification method termed BioTAP-XL to preserve cMYC integrity and chromatin association for genomic and mass spectrometry-based proteomic analyses. We found that in the NC797 cell line, cMYC interacts primarily with the NuA4 KAT5 lysine acetyltransferase complex, with interactions that are maintained despite a decrease in cMYC levels after JQ1 treatment. We propose that a cascade of aberrant acetyltransferase activities in NC797 cells, first via EP300 recruitment by B4N to mis-regulate *cMYC*, and then by KAT5 interaction with the resulting overexpressed cMYC protein, drive NC cell proliferation and blockade to differentiation.

Keywords: NUT carcinoma; cMYC; NuA4 acetyltransferase; KAT5; BioTAP-XL

44 **Graphical Abstract:**

45



46

47

48

49 **1. Introduction**:

50 NUT-rearrangement cancers are a collection of aggressive and poorly-differentiated squamous cell
51 carcinomas whose unique molecular hallmark is somatic expression of *NUTM1*
52 (HGNC:29919)[1]. Chromosomal rearrangements result in the otherwise germline-restricted
53 *NUTM1* gene to be fused downstream of a somatically expressed gene, most commonly *BRD4*
54 (HGNC:13575) to generate the fusion protein BRD4-NUT (B4N). At least two NC fusion proteins:
55 B4N and ZNF532-NUT (a fusion product with *ZNF532* (HGNC:30940)) have been shown to
56 regulate expression of the proto-oncogene *cMYC* (hereafter referred to as *MYC*). This is thought
57 to occur through an aberrant, feed-forward loop in which the NUTM1 portion of the fusion protein
58 attracts the EP300 lysine acetyltransferase, while the BRD4 bromodomain binds the resulting
59 acetylated nucleosomes. The resulting formation of unusually large, hyperacetylated chromatin
60 regions termed "megadomains" drives increased expression of underlying genes and regulatory
61 regions [2, 3]. Treatment of NC patient cells in culture with a small molecule, JQ1, that inhibits
62 interaction of the BRD4 bromodomains with acetylated lysines, leads to loss of megadomains,
63 down-regulation of *MYC,* and differentiation of B4N cells [2, 4]. The unique importance of *MYC*
64 was revealed in studies where induced *MYC* transgene expression reversed the differentiation
65 induction caused by B4N knockdown [5]. Furthermore, non-B4N-dependent pathways such as
66 *KLF4* (HGNC:6348) expression help maintain *MYC* levels in cells resistant to JQ1 [6]. Thus,
67 understanding the mechanism of action of MYC in NC cells should provide valuable insights into
68 this therapeutically intractable disease.

69

70 Much of what is understood about MYC stems from non-NC cells [7-13]. MYC binds DNA as a
71 heterodimer with MAX, a second bHLHZip (basic helix-loop-helix leucine zipper) protein [14]

72 and positively regulates the expression of critical enzymes involved in growth and metabolism,
73 thereby increasing global RNA and protein expression [9]. Although the direct target specificity
74 of MYC has been a subject of considerable debate, this is likely due to an incomplete picture of
75 the relationship between binding and function. MYC can be observed to bind to most transcription
76 start sites (TSS), enhancers, and accessible chromatin in various cell types, as it has relatively
77 strong non-specific DNA binding in addition to its sequence-specificity for E-box motifs [15, 16].
78 Current genetic and functional analyses are consistent with direct regulation by MYC of genes
79 involved in nucleotide biosynthesis, RNA metabolism, and ribosome biogenesis, which in turn
80 drive numerous secondary changes in gene expression [10, 16].
81
82 In addition to heterodimerization with MAX, early studies identified MYC protein interactions
83 with the SAGA/STAGA and NuA4 histone acetyltransferase complexes [17-22], linked to MYC
84 via the TRRAP adaptor protein [23, 24]. These multi-subunit complexes [25, 26] each function
85 as transcriptional coactivators, with enzymatic activities targeted mainly to histone H3 in the case
86 of SAGA/ STAGA, and H2A, H2AZ, and H4 in the case of NuA4 [24].
87
88 For many years comprehensive proteomic analyses of MYC interactions were impeded by
89 technical challenges. First, in normal cells the MYC protein has a relatively short half-life of
90 approximately 30 minutes, a result of ubiquitin-dependent and ubiquitin-independent pathways
91 and of the even shorter (10 minute) half-life of *MYC* RNA itself [27-30]. Second, expression of
92 the *MYC* gene depends on many long-range distal regulatory elements [31-34] suggesting the
93 importance of complex regulatory circuits of the *MYC* locus which ideally should be preserved
94 when analyzing MYC in different cell states. Finally, it has been difficult to extract intact MYC

complexes from chromatin using traditional biochemical methods. This technical challenge has been overcome by proximity-dependent biotin identification (BioID) with considerable success [12, 35]. Here, we apply an orthogonal approach, BioTAP-XL [3, 36], to characterize MYC protein and genomic interactions in NC and non-NC cells. BioTAP-XL employs crosslinking, stringent two-step affinity purification, and analysis of enrichment over input, thus diminishing background and potentially allowing the most significant interactions to rise to the top. Consistent with the previous studies [12, 17-22, 35, 37], we found that subunits of the NuA4 and STAGA lysine acetytransferases were top MYC interactors in a non-NC cell line. In contrast, NuA4 was the predominant MYC interactor in an NC cell line, and these interactions were largely maintained after JQ1 treatment reduced overall MYC levels, suggesting that protein complex assembly may be independent of differentiation status. As chromatin acetylation complexes were the most enriched MYC interactors using our stringent cross-linking, tandem affinity purification approach, our results support models in which acetylation is likely to be central to the biological function of MYC.

111 **2. Materials and Methods:**

112 *2.1. Generation of N-BioTAP-cMYC NC797 cells using CRISPR/Cas9*

113 To generate N-BioTAP and N-ProteinA (PrA) cMYC fusions in NC797 cells [38], the following
114 DNA fragments: pAVV-MCS-NheI-5'cMyc_BamHI-ATG-loxP-Blasti-P2A-loxP-PrA-Bio-
115 3'cMyc and pAAV-5'-cMYC-ATG-LoxP-Blasti-P2A-loxP-PrA-3'-cMYC respectively, were
116 synthesized as gBlock gene fragments (Integrated DNA Technologies, **Datafile S1**) and
117 introduced into the NheI/BamHI restriction sites of pAAV-MCS2 by Gibson assembly (New
118 England Biolabs). N-BioTAP and N-PrA cassettes targeted the first ATG of the *MYC* gene. The
119 pAAV-nEFCas9 (Addgene, plasmid 87115) was used to express Cas9, and the pAAV-tagBFP U6-
120 gRNA expression vector was used to express a gRNA targeting the start codon of cMYC: gRNA-
121 cMYC-start (**Table 1**). Recombinant AAV2 was packaged in HEK293T cells using pHelper and
122 pRC2-mi342 plasmids (Takara, Catalog #632608). Three days after transfection, cells were
123 harvested and AAV2 was isolated using AAVpro Extraction Solution (Takara, Catalog #6235).
124 Two million NC797 cells were infected with 30ul of each adeno-associated virus (AAV2).
125 Following 1 week of culture post-AAV infection, cells were selected with 7.5ug/ml Blasticidin
126 (Invitrogen**).** Surviving cells were re-plated and single-cell clones were isolated and expanded, and
127 genotyping PCR and sequencing were performed to check for proper integration of the BioTAP or
128 PrA cassette in the c-MYC locus. We recovered both heterozygous and homozygous clones.
129 Homozygosity of the cassette insertion was determined by absence of the wildtype PCR fragment
130 (1033bp) and presence of the larger PCR fragment (2083bp) with the following PCR primers:
131 Genom-Myc(1S) and Genom-Myc(1A) (**Table 1**). Next, we validated the presence of the correctly
132 sized tagged MYC protein and absence of wild type MYC by Western blot probed with a c-MYC
133 antibody (Rabbit mAb, Cell Signaling Technology, Cat# 5605). We selected homozygous clone

134 number 2 (for N-BioTAP-cMYC) and clone number 4 (for N-PrA-cMYC) for Cre-dependent
135 cassette excision experiments.

136 **Table 1. Sequences of oligonucleotides used in paper.**

Name	Sequence (5′ to 3′)
gRNA-cMYC-start	GACGTTGAGGGGCATCGTCGC
Genom-Myc(1S)	GAGTGGGAACAGCCGCAG
Genom_Myc(1A)	CAGGTACAAGCTGGAGGTGG

137

138 Next, the recombinant pAAV-CRE plasmid was packaged in HEK293T cells using pHelper and
139 pRC2-mi342 plasmids (Takara, Catalog #632608). Cells were harvested three days after
140 transfection and AAV2 was isolated using AAVpro Extraction Solution (Takara, Catalog #6235).
141 Two million 797 cells with homozygous insertions (clones 2 or 4) were infected with 30ul of each
142 adeno-associated CRE virus (AAV2). Following one week of culture post-AAV infection, cells
143 were re-plated, single-cell clones were isolated and expanded, and genotyping PCR and
144 sequencing were performed to check for absence of the drug resistance cassette in the c-MYC
145 locus. We selected clone number 2-7 (for N-BioTAP-cMYC) and clone number 4-6 (for N-PrA-
146 cMYC) to proceed with LC-MS and ChIP-seq experiments.

147 *2.2. Generation of inducible N and C-BioTAP-cMYC expressing HEK293T cells*

148 We used lentiviral transduction to generate stable HEK293T cell lines [2] carrying cMYC
149 constructs with the BioTAP tag at either the N or the C terminus to induce transcription of these
150 clones from the CMV/2xtetO promoter for genetic complementation experiments. To make N- and
151 C-terminal BioTAP constructs, we used the Gateway recombination system to introduce the
152 following Myc ORF clone HsCD00039771 in pDONR221from the CCSB Human ORFeome
153 Collection into pHAGE-TRE-DEST-NBioTAP (Addgene #53568) and pHAGE-TRE-DEST-
154 CBioTAP lentiviral vectors (Addgene #53569). Stable 293T-Trex-cell lines containing each

155 transgene were generated by lentiviral transduction in the presence of 8μg/mL polybrene, followed
156 by selection with 2μg/mL puromycin dihydrochloride (Sigma #P8833) as described previously
157 [39]. To induce transcription of the cDNA clones from the CMV/2xtetO promoter we added
158 doxycycline (1μg/mL) to the medium for 48 hours.

159 *2.3. Proteomic analysis of cMYC interactions using BioTAP-XL*

160 All cell lines were cultured as monolayers in DMEM (Invitrogen) supplemented with 1x Penicillin
161 Streptomycin (Hyclone, South Logan, UT), 1x Glutamax (Gibco), and 10% (v/v) Fetal bovine
162 serum (FBS) (Hyclone). For JQ1 treatment, 500uM of JQ1 (gift from Dr. C.A. French) was added
163 to NC797 cells for various durations before harvesting for Western Blot analysis. The 12 hour
164 timepoint was ultimately chosen for BioTAP-XL experiments.
165 Our primary experiments were performed with 2-7 N-BioTAP-cMYC (clone 2-7). 2.0 × 10^9 cells
166 were grown as monolayer cultures in 150-mm dishes. 100 plates total were used for BioTAP-XL
167 purification. The main steps of the BioTAP-XL procedure were as previously described [39]. For
168 single-step BioTAP-XL, proteins enriched by IgG-agarose beads were eluted and processed with
169 Agencourt AMPure-XP magnetic beads (Beckman Coulter) as previously described [40]. In brief,
170 fifty micrograms of eluted proteins were acidified and diluted with acetonitrile prior to binding by
171 magnetic beads. SDS was removed with ethanol and acetonitrile washes, and proteins were
172 digested with overnight trypsin incubation, yielding peptides for desalting and LC-MS analysis.
173 LC-MS acquisition and analyses were performed as previously described [3]. Offline peptide
174 fractionation was not performed. Rather, peptide samples were autosampler loaded onto a 100μm
175 ID fused silica capillary column packed with 25cm of 2.6μm Accucore C18-resin (Thermo Fisher)
176 and electrosprayed into either a Velos-Orbitrap Pro or Velos-Orbitrap Elite mass spectrometer
177 (Thermo Fisher). Peptides were resolved with a gradient from 0% to 5% Buffer B from 0 to 10

178 minutes, from 5% to 32% Buffer B from 10 to 65 minutes, and from 32% to 95% Buffer B from
179 65 to 68 minutes at a flow rate of approximately 110nl/min delivered on an Accela pump (Thermo
180 Fisher), where Buffer A composition is 97.5% water, 2.4% acetonitrile, 0.1% formic acid and
181 Buffer B composition is 7.4% water, 92.5% acetonitrile, 0.1% formic acid. Peptides were analyzed
182 under Top22 data-dependent mode with collision-induced dissociation (CID), using the Orbitrap
183 detector for MS1 acquisition and ion-trap detector for MS2 acquisition. Peptides were identified
184 with SEQUEST against the human proteome with precursor mass tolerance at 20ppm and fragment
185 ion mass tolerance at 0.9 Da. Methionine oxidation was set as dynamic modification.
186 Identifications were further filtered at a 2% false discovery rate. The number of total spectra
187 number for each protein were used to determine relative abundance. Up to two technical replicates
188 per input or immunoprecipitated samples were acquired.

189 *2.3. Genomic analysis of cMYC binding*

190 ChIP-seq analysis was performed in parallel with proteomic analyses as previously described [36].
191 Raw fastq files were aligned to the human genome (hg38) using bowtie2.3.4.3 with default settings
192 [41]. Duplicate reads were then removed with samtools1.14 [42]. Peak calling was performed on
193 DMSO treated samples with HOMER4.9 using the findPeaks command [43]. HOMER
194 annotatePeaks.pl was used to assign peaks to genomic regions. Deeptools3.0.2 was used to
195 generate bigwig files, enrichment profiles, and heatmaps [44]. Bigwig files were normalized to 1x
196 genome coverage. For heatmaps and enrichment profiles, MYC signal was normalized to its
197 corresponding input and enrichment was shown as $\log_2$(IP/input). Differential peak analysis was
198 performed with the DiffBind package in RStudio using default settings:
199 (http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf).Dataset
200 s can be obtained from the NCBI GEO database (accession# GSE201033).

201 **3. Results:**

202 *3.1. BioTAP-XL strategy for analysis of MYC interactions*

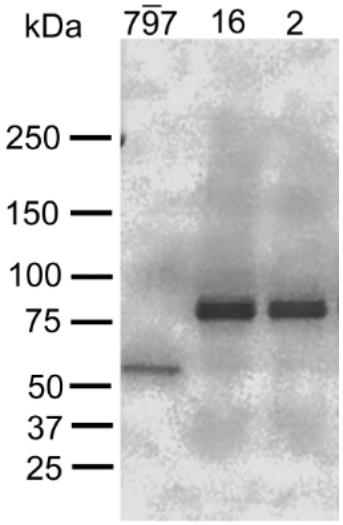
203 To address the potential instability or loss of MYC during conventional biochemical extraction
204 from DNA, we applied a crosslinking approach termed BioTAP-XL for MYC complex
205 purification in NC797 cells, an NC patient-derived cell line, and in HEK293 cells, a non-NC line
206 for comparison. The BioTAP affinity tag itself is composed of the Z domain from *Staphylococcal*
207 protein A and a biotinylation target domain from *Propionibacterium* transcarboxylase.
208 Biotinylation of this target sequence occurs in most organisms without requiring introduction of
209 an exogenous biotin ligase or addition of biotin to the media [36]. After generating a cell line that
210 expresses the BioTAP-tagged bait, the BioTAP-XL workflow involves three steps: 1)
211 formaldehyde to inactivate cellular enzymes and to crosslink proximal protein and nucleic acid
212 interactions, 2) sonication to solubilize the crosslinked chromatin pellet, and 3) enrichment of the
213 tagged bait using tandem ProteinA and streptavidin pulldowns combined with stringent washes to
214 reduce nonspecific interactions. This approach has successfully characterized the proteomic and
215 genomic interactions of B4N, wild-type BRD4, ZNF532-NUT, wild-type ZNF532, and other
216 proteins [3].

217

218 *3.2. NuA4 lysine acetyltransferase subunits are highly enriched in BioTAP-MYC purifications*

219 To characterize the protein interactions of N-BioTAP-MYC in NC797 cells, we used a CRISPR-
220 Cas9 mediated knock-in approach. We successfully inserted the BioTAP affinity tag into both
221 copies of the endogenous *cMYC* gene in NC797 cells (see Methods, **Figure 1**). A critical concern
222 with interpreting affinity tag results is whether the tagged protein preserves the functions of the
223 non-tagged protein. In NC models such as NC797 cells, B4N regulates transcription of the *MYC*

224 gene, and NC797 cells are dependent on functional MYC expression [1, 2, 4, 5]. Since knockdown
225 of MYC was previously shown to induce differentiation of NMC cells [5], that we could create
226 and grow homozygous knock-in NC797 cells in large numbers (>10^9) strongly suggests that the
227 epitope tags do not interfere with normal MYC function.
228



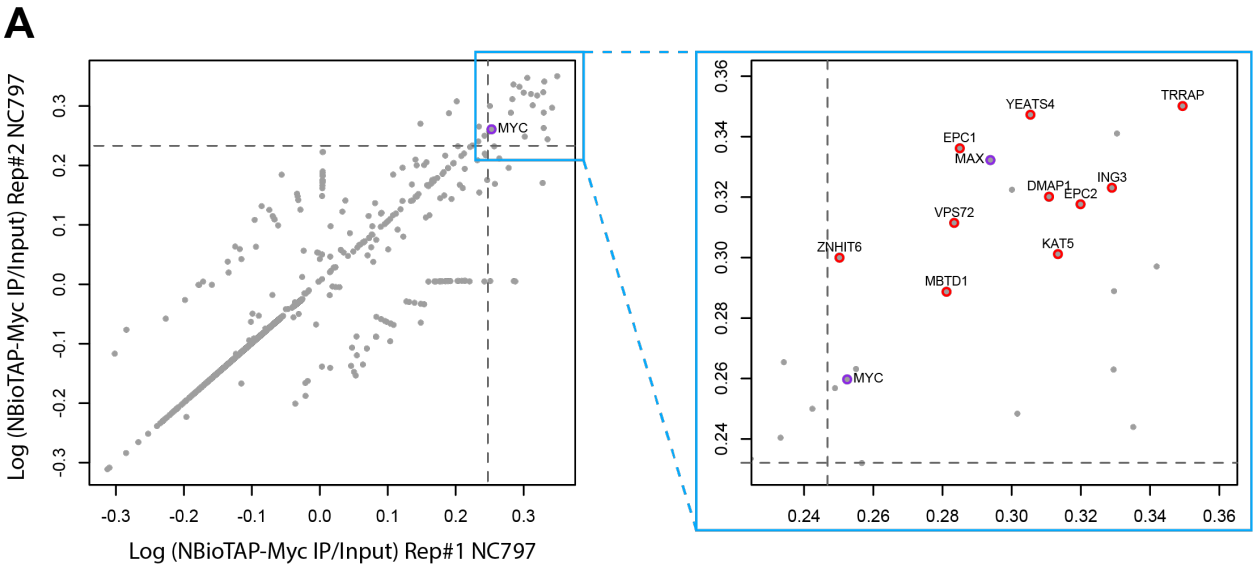
229

230 **Figure 1.** Western blot of NC797 cells carrying homozygous knock-in alleles of BioTAP-tagged
231 MYC probed with anti-MYC (D84C12 Rabbit mAb, Cell Signalling Technologies). The untagged
232 size of MYC in unedited NC797 cells is ~57 Kd, (lane 1). The size of N-BioTAP-MYC is ~80
233 Kd, seen in Clones 16 and 2 (lanes 2 and 3). Clone 2 was used for subsequent experiments. The
234 original gel image can be viewed in Supplemental Figure 1.

235

236 We performed BioTAP-XL to enrich for MYC containing protein complexes and used liquid
237 chromatography-mass spectrometry (LC-MS) to sequence the peptides resulting from trypsin
238 digestion of the enriched complexes. MYC protein interactions were identified as the proteins with
239 the greatest fold enrichment in total peptides identified in the pulldown compared to the input

240 sample (**Datafile S2**). Fold-enrichment was calculated as the number of total peptides mapping to
241 each protein divided by the molecular weight of the protein as a mass-normalized ratio, which was
242 subsequently divided by the sum of all the mass-normalized ratios across all proteins identified in
243 a given LCMS sample, and finally, compared between samples as log ratios.
244
245 We observed consistency (Pearson > 0.8) in overall enrichment of proteins between technical
246 replicates of our BioTAP-XL experiments (**Supplemental Figure 2**). As expected, we recovered
247 both MYC and its heterodimerization partner MAX [14] as among the top one to two percentile
248 most enriched proteins (corresponding in this case to approximately twenty-three proteins).
249 Beyond MYC and MAX, eleven of the most enriched proteins are related to chromatin function.
250 In particular, eight are components of the NuA4 histone acetyltransferase complex [26, 37]
251 (**Figure 2**). The enriched components include DMAP1 (HGNC:18291), EP400 (HGNC:11958),
252 EPC1 (HGNC:19876), EPC2 (HGNC:24543), KAT5 (HGNC:5275), MBTD1 (HGNC:19866),
253 TRRAP (HGNC:12347), and YEATS4 (HGNC:24859).
254



B

				MYC BioTAP		Myc PrA Vehicle		Myc PrA JQ1		
Name	MolWt	Input	JQ1 Input	Rep #1	Rep #2	Rep #1	Rep #2	Rep #1	Rep #2	No Bait IgG
MYC	50.53	0	0	3	3	6	8	3	6	0
MAX	18.26	0	0	3	4	7	9	6	7	0
DMAP1	52.96	1	2	9	9	9	8	5	6	0
EP400	343.28	8	7	58	66	53	47	34	30	0
EPC1	93.41	0	0	6	10	4	9	2	4	0
EPC2	91.04	0	0	9	8	9	9	6	8	0
ING3	46.71	0	0	7	6	6	7	5	4	0
KAT5	61.76	1	0	10	8	4	6	1	3	0
MBTD1	70.5	0	0	5	5	3	2	2	3	0
TRRAP	437.32	8	8	107	98	88	85	52	79	0
VPS72	41.82	0	1	4	5	4	4	2	3	0
YEATS4	26.48	1	0	6	8	2	3	3	2	0
ZNHIT6	53.88	0	0	3	5	1	1	0	1	0

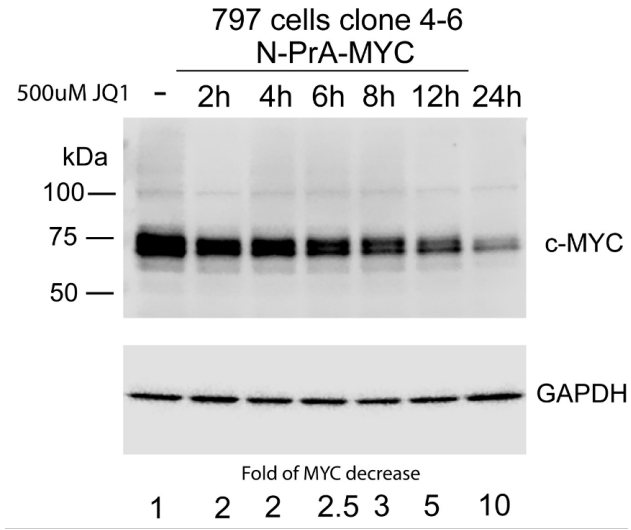
Figure 2. Characterization of MYC protein interactions in NC797 cells. (a) Scatterplots depicting enrichment of proteins from N-terminally tagged MYC using BioTAP-XL workflow in NC797 cells compared to non-enriched input. Each gray dot represents an individually identified protein, where its position is a measure of enrichment efficiency based on the number of total peptides identified in pulldown compared to input and normalized to the molecular weight of the protein from two technical replicates. The dashed gray lines denote the 98th percentile of enrichment. The right plot is a zoomed in region of the most enriched proteins in the left plot from replicate experiments. (b) Number of total peptides

264 mapping to each protein from a given pulldown experiment in independent (BioTAP- vs. PrA-
265 tagged) NC797 cell isolates.

266

267 *3.3. Comparison of MYC interactions in distinct cell types and cell states*

268 To test whether the enrichment of NuA4 but not STAGA might be related to our crosslinking
269 method rather than the NC cell type, we performed the BioTAP-XL experiment in the non-NC cell
270 line HEK293-TREx, which was used in previous studies [12]. Similar to NC797 cells, we found
271 enrichment of NuA4 components with the MYC-BioTAP bait in HEK293-TREx cells, but we also
272 observed enrichment of SAGA/STAGA components such as ATXN7 (HGNC:10560) and KAT2A
273 (HGNC:4201) (**Supplemental Figure 3, Datafile S3**). Therefore, MYC interaction with NuA4
274 but not STAGA appears to be a characteristic of the NC797 cell line rather than somehow biased
275 by our crosslinking approach.



283 **Figure 3.** Effect of JQ1 treatment on MYC protein expression. Levels of MYC protein in N-PrA-cMyc-NC797 cells were assayed by Western blot, following different durations of 500uM JQ1 treatment, with approximate fold change in MYC levels provided below the blot compared to no treatment. The original gel image can be seen in Supplemental Figure S1.

284

285 To test whether the NuA4 interaction with MYC would change upon B4N inactivation and the
286 resulting cell state transition towards differentiation, we performed affinity enrichment of MYC
287 after cells were treated with the bromodomain inhibitor JQ1. We knew from prior experience that
288 the amount of MYC bait protein expression after inactivation of B4N would be significantly
289 decreased, which might be insufficient for a full tandem affinity purification. Therefore we
290 modified our approach to utilize a single rather than dual tag, creating NC797 clones in which the
291 endogenous gene had only the ProteinA portion of the affinity tag (PrA-MYC). Consistent with
292 B4N directly regulating the endogenous *MYC* locus, treatment of NC797 cells with JQ1 reduced
293 levels of PrA-MYC in a time-dependent manner (**Figure 3**). That cells relying on epitope-tagged-
294 MYC still responded to JQ1 provided further evidence for *MYC* functionality after tag insertion.
295 Although we had hoped to uncover regulation of MYC at the level of protein-protein interactions,
296 we found that NuA4 interactions with MYC were largely preserved after a 12 hour treatment with
297 JQ1 corresponding to a five-fold reduction in PrA-MYC levels (**Figure 2B**). To test whether
298 recovery of NuA4 components could be due to non-specific capture of the proteins by the IgG
299 agarose beads, we performed a parallel pulldown with NC797 cells expressing no tagged bait. We
300 observed no recovery of peptides from MYC or NuA4 components (**Figure 2B**). Taken together,
301 our results confirm the strong relationship between MYC and NuA4 in NC797 cells.
302

303 *3.4. Genomic profiling confirms broad MYC association largely unchanged by JQ1 treatment*

304 In parallel with the proteomic analyses, we mapped the genomic occupancy of PrA-MYC in
305 NC797 cells following JQ1 or DMSO treatment. Differential binding analysis resulted in 2,238
306 regions with significantly altered MYC binding (FDR < 0.05) (**Figure 4A,B**). Consistent with the
307 decrease in MYC levels, all but 4 of these MYC-bound regions showed decreased occupancy

308 (n=2,234). Significantly downregulated regions were located at promoters, defined as +/- 1 Kb
309 from transcription start sites (25%, n=555/2,234), introns (40%, n=900/2,234), and intergenic
310 regions (28%, n=623/2,234) (**Figure 4C)**. While MYC occupancy was diminished, the identities
311 of MYC binding sites were not globally changed by JQ1 treatment (**Figure 4B,D**). It is known that
312 JQ1 treatment alters gene expression [1, 2], but in retrospect, the lack of overall change in genomic
313 occupancy is not surprising given the recent discovery that pervasive non-specific DNA binding
314 contributes to, but also obscures the considerably smaller set of specific regulatory targets of MYC
315 [15, 16]. In summary, our comparison of proliferating and differentiating NC cells suggests that
316 MYC-dependent changes in cell state generally occur without altering the identity of the most
317 enriched MYC protein interactions or overall DNA binding sites.
318

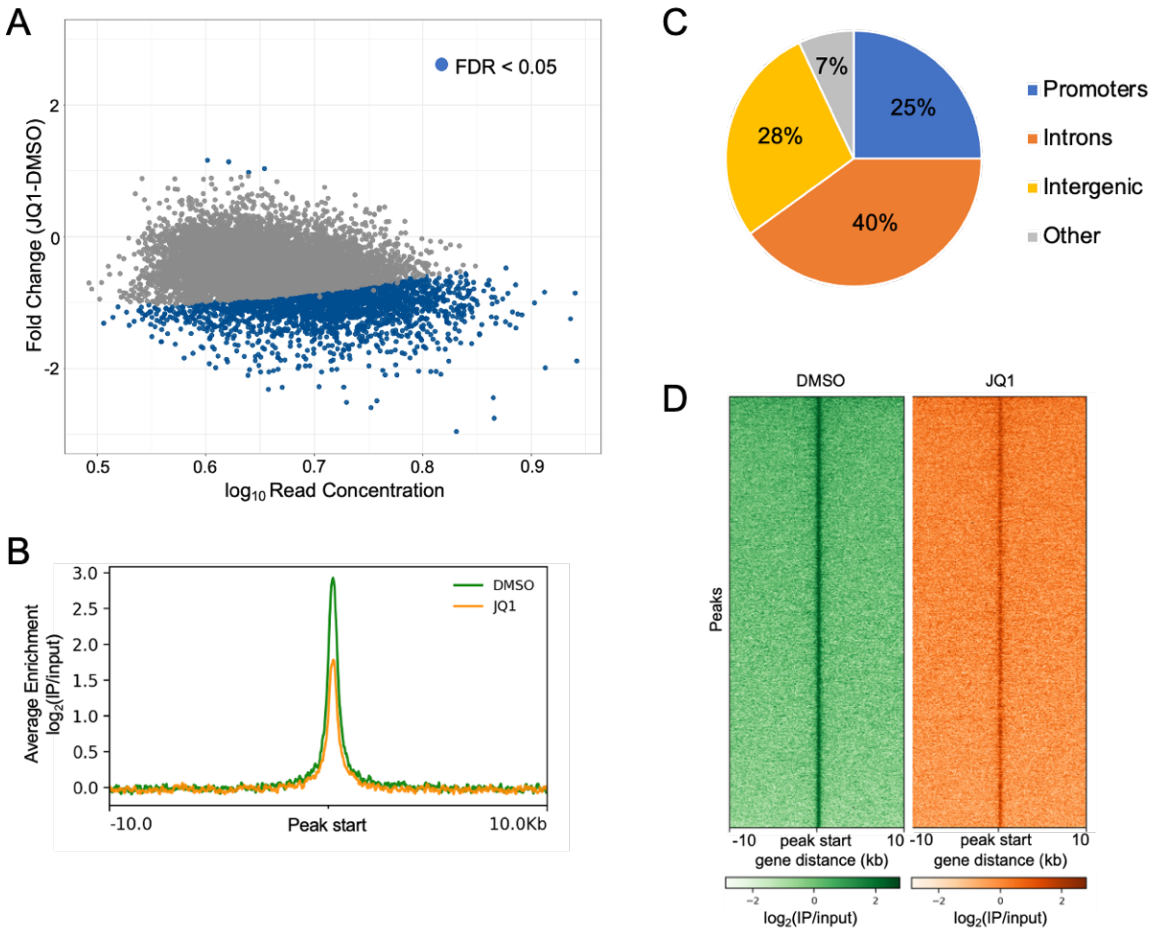


Figure 4. Effect of JQ1 on MYC genomic binding. (a.) MA plot of MYC differential binding after JQ1 treatment. Each point represents a 400 bp MYC-bound region. Blue points indicate significantly altered regions (FDR < 0.05). The x-axis depicts the $\log_{10}$ of the average read concentration across all samples. The y-axis indicates the fold change after JQ1 treatment. (b.) Enrichment profile plotting the average signal of significantly decreased MYC-bound regions in DMSO and JQ1 treated samples. One representative replicate is shown. (c.) Pie chart depicting the percentage of significantly decreased MYC-bound regions mapped to the specified genomic regions. (d.) Heatmap of MYC enrichment at significantly decreased MYC-bound regions in DMSO and JQ1 treated samples. One representative replicate is shown.

329 **4. Discussion:**

330 To characterize MYC protein partners in NC797 cells, we applied a crosslinking affinity pulldown
331 method. We took this approach because, historically, MYC complexes have been difficult to
332 purify. However, significant progress has been made in this regard through recent BioID studies
333 by the Penn laboratory, in which proximity labeling resulted in a comprehensive study of MYC
334 interactors [12]. Extending the known association with acetyltransferases [17-22], they mapped
335 the interactions of both NuA4 and STAGA complexes to MYC homology box II (MBII), showed
336 that MBII is linked to both complexes through TRRAP, and confirmed the importance of MBII in
337 stimulation of acetylation by MYC. Our data in HEK293T cells using the orthogonal BioTAP-XL
338 approach is in agreement with the association of MYC with both NuA4 and STAGA. It remains
339 to be determined why the NC cells we studied here show strong enrichment for NuA4 but not
340 STAGA. In this regard, we hope that our datasets will be useful in further study of MYC and
341 downstream targets specifically in NC cells. Demonstrations that MYC is dependent on
342 acetyltransferases for its function have suggested potential therapeutic options to what is generally
343 considered an undruggable oncoprotein [45, 46]. Our results suggest that development of specific
344 inhibitors of KAT5 for synergy with JQ1 could be particularly relevant to the treatment of NC.
345

346 **5. Conclusion**

347 MYC is a central player in a majority of cancers, yet understanding its regulation, targeting, and
348 mode of action are important unsolved problems. When we initiated this study, we knew that the
349 B4N oncogenic fusion attracted the p300 acetyltransferase to drive oncogenesis, with the *MYC*
350 gene as a critical target [1]. Since MYC function in NC cells is reversible by the JQ1 small
351 molecule BET bromodomain inhibitor, we saw this as an opportunity to compare MYC

352 interactions in the same cells under oncogenic and differentiating conditions. While we did not
353 find a difference in protein interactions or gene targeting in the distinct growth states, we identified
354 a close association between MYC and the NuA4 acetyltransferase in NC cells. We propose that
355 future analysis of MYC transcriptional regulation in NUT carcinoma, and likely other cancers,
356 should prioritize how MYC recruits KAT complexes to modify the chromatin of its many gene
357 targets.
358
359 **Supplementary Materials:**
360 Fig. S1: Original gel images for Figs. 1 and 3.
361 Fig. S2. Correlation of NC797 BioTAP-XL replicates.
362 Fig. S3. Comparison of MYC interactions with NuA4 and SAGA/STAGA components in HEK293
363 and NC797 cells.
364 Datafile S1: DNA sequences used to create NC797 cells with N-terminal fusions of the BioTAP-
365 or Protein A-epitope tags to the endogenous cMYC coding region.
366 Datafile S2: Total peptide counts of NC797 proteomic results.
367 Datafile S3: Total peptide counts of HEK293T proteomic results.
368
369 **Author Contributions:**
370 Conceptualization, AAA and MIK; methodology, validation, formal analysis, investigation, BMZ
371 and AAA; resources, data curation, BMZ, AES, and HJK; writing – original draft preparation BMZ
372 and AAA; writing – review and editing, BMZ and MIK; visualization, BMZ, AAA, AES and HJK;
373 supervision, project administration, funding acquisition, MIK.
374

Funding:

This research was funded by the National Institutes of Health, grant number R35 GM126944 to MIK. AES is supported by an American Cancer Society postdoctoral fellowship 721690.

Data Availability Statement

Proteomic data sets can be found at Harvard Dataverse (https://doi.org/10.7910/DVN/IFKRBT). Genomic data sets can be found at NCBI GEO database (accession# GSE201033).

Acknowledgments

We thank Dr. C.A. French for introducing us to NC research, and Dr. J.L. Makofske for initial work on this project. We thank Ross Tomaino (Taplin Mass Spectrometry Facility, Harvard Medical School) for expert help with analyzing the proteomics samples. We thank Drs. T. Martin and S.J. Elledge for the pAAV-tagBFP U6-gRNA expression vector. Finally, we thank H.J. Kang for helpful assistance with data submission.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. French, C.A., *NUT Carcinoma: Clinicopathologic features, pathogenesis, and treatment.* Pathol Int, 2018. **68**(11): p. 583-595.
2. Alekseyenko, A.A., et al., *The oncogenic BRD4-NUT chromatin regulator drives aberrant transcription within large topological domains.* Genes & development, 2015. **29**(14): p. 1507-23.
3. Alekseyenko, A.A., et al., *Ectopic protein interactions within BRD4-chromatin complexes drive oncogenic megadomain formation in NUT midline carcinoma.* Proc Natl Acad Sci U S A, 2017. **114**(21): p. E4184-E4192.
4. Filippakopoulos, P., et al., *Selective inhibition of BET bromodomains.* Nature, 2010. **468**(7327): p. 1067-73.
5. Grayson, A.R., et al., *MYC, a downstream target of BRD-NUT, is necessary and sufficient for the blockade of differentiation in NUT midline carcinoma.* Oncogene, 2014. **33**(13): p. 1736-1742.
6. Liao, S., et al., *Genetic modifiers of the BRD4-NUT dependency of NUT midline carcinoma uncovers a synergism between BETis and CDK4/6is.* Genes Dev, 2018. **32**(17-18): p. 1188-1200.

411 7. Meyer, N. and L.Z. Penn, *Reflecting on 25 years with MYC.* Nat Rev Cancer, 2008. **8**(12):
412 p. 976-90.
413 8. Dang, C.V., *MYC on the path to cancer.* Cell, 2012. **149**(1): p. 22-35.
414 9. Kress, T.R., A. Sabo, and B. Amati, *MYC: connecting selective transcriptional control to*
415 *global RNA production.* Nat Rev Cancer, 2015. **15**(10): p. 593-607.
416 10. Muhar, M., et al., *SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC*
417 *axis.* Science, 2018. **360**(6390): p. 800-805.
418 11. Carroll, P.A., et al., *The MYC transcription factor network: balancing metabolism,*
419 *proliferation and oncogenesis.* Front Med, 2018. **12**(4): p. 412-425.
420 12. Kalkat, M., et al., *MYC Protein Interactome Profiling Reveals Functionally Distinct*
421 *Regions that Cooperate to Drive Tumorigenesis.* Mol Cell, 2018. **72**(5): p. 836-848 e7.
422 13. Lourenco, C., et al., *MYC protein interactors in gene transcription and cancer.* Nat Rev
423 Cancer, 2021. **21**(9): p. 579-591.
424 14. Blackwood, E.M. and R.N. Eisenman, *Max: a helix-loop-helix zipper protein that forms a*
425 *sequence-specific DNA-binding complex with Myc.* Science, 1991. **251**(4998): p. 1211-7.
426 15. Guo, J., et al., *Sequence specificity incompletely defines the genome-wide occupancy of*
427 *Myc.* Genome Biol, 2014. **15**(10): p. 482.
428 16. Pellanda, P., et al., *Integrated requirement of non-specific and sequence-specific DNA*
429 *binding in Myc-driven transcription.* EMBO J, 2021. **40**(10): p. e105464.
430 17. McMahon, S.B., M.A. Wood, and M.D. Cole, *The essential cofactor TRRAP recruits the*
431 *histone acetyltransferase hGCN5 to c-Myc.* Mol Cell Biol, 2000. **20**(2): p. 556-62.
432 18. Bouchard, C., et al., *Regulation of cyclin D2 gene expression by the Myc/Max/Mad*
433 *network: Myc-dependent TRRAP recruitment and histone acetylation at the cyclin D2*
434 *promoter.* Genes Dev, 2001. **15**(16): p. 2042-7.
435 19. Frank, S.R., et al., *MYC recruits the TIP60 histone acetyltransferase complex to chromatin.*
436 EMBO Rep, 2003. **4**(6): p. 575-80.
437 20. Frank, S.R., et al., *Binding of c-Myc to chromatin mediates mitogen-induced acetylation of*
438 *histone H4 and gene activation.* Genes Dev, 2001. **15**(16): p. 2069-82.
439 21. Fernandez, P.C., et al., *Genomic targets of the human c-Myc protein.* Genes Dev, 2003.
440 **17**(9): p. 1115-29.
441 22. Knoepfler, P.S., et al., *Myc influences global chromatin structure.* EMBO J, 2006. **25**(12):
442 p. 2723-34.
443 23. McMahon, S.B., et al., *The novel ATM-related protein TRRAP is an essential cofactor for*
444 *the c-Myc and E2F oncoproteins.* Cell, 1998. **94**(3): p. 363-74.
445 24. Cheung, A.C.M. and L.M. Diaz-Santin, *Share and share alike: the role of Tra1 from the*
446 *SAGA and NuA4 coactivator complexes.* Transcription, 2019. **10**(1): p. 37-43.
447 25. Chen, Y.C. and S.Y.R. Dent, *Conservation and diversity of the eukaryotic SAGA*
448 *coactivator complex across kingdoms.* Epigenetics Chromatin, 2021. **14**(1): p. 26.
449 26. Doyon, Y. and J. Cote, *The highly conserved and multifunctional NuA4 HAT complex.* Curr
450 Opin Genet Dev, 2004. **14**(2): p. 147-54.
451 27. Hann, S.R. and R.N. Eisenman, *Proteins encoded by the human c-myc oncogene:*
452 *differential expression in neoplastic cells.* Mol Cell Biol, 1984. **4**(11): p. 2486-97.
453 28. Hickman, K.A., et al., *Image-Based Analysis of Protein Stability.* Cytometry A, 2020.
454 **97**(4): p. 363-377.
455 29. Murai, N., et al., *Novel ubiquitin-independent nucleolar c-Myc degradation pathway*
456 *mediated by antizyme 2.* Sci Rep, 2018. **8**(1): p. 3005.

30. Dani, C., et al., *Extreme instability of myc mRNA in normal and transformed human cells.* Proc Natl Acad Sci U S A, 1984. **81**(22): p. 7046-50.
31. Cole, M.D., *MYC association with cancer risk and a new model of MYC-mediated repression.* Cold Spring Harb Perspect Med, 2014. **4**(7): p. a014316.
32. Cho, S.W., et al., *Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element.* Cell, 2018. **173**(6): p. 1398-1412 e22.
33. Fulco, C.P., et al., *Systematic mapping of functional enhancer-promoter connections with CRISPR interference.* Science, 2016. **354**(6313): p. 769-773.
34. Tseng, Y.Y., et al., *PVT1 dependence in cancer with MYC copy-number increase.* Nature, 2014. **512**(7512): p. 82-6.
35. Dingar, D., et al., *BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors.* J Proteomics, 2015. **118**: p. 95-111.
36. Alekseyenko, A.A., et al., *BioTAP-XL: Cross-linking/Tandem Affinity Purification to Study DNA Targets, RNA, and Protein Components of Chromatin-Associated Complexes.* Curr Protoc Mol Biol, 2015. **109**: p. 21 30 1-21 30 32.
37. Kim, J., et al., *A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs.* Cell, 2010. **143**(2): p. 313-24.
38. Toretsky, J.A., et al., *Translocation (11;15;19): a highly specific chromosome rearrangement associated with poorly differentiated thymic carcinoma in young patients.* Am J Clin Oncol, 2003. **26**(3): p. 300-6.
39. Alekseyenko, A.A., et al., *Reciprocal interactions of human C10orf12 and C17orf96 with PRC2 revealed by BioTAP-XL cross-linking and affinity purification.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(7): p. 2488-93.
40. Zee, B.M., et al., *Streamlined discovery of cross-linked chromatin complexes and associated histone modifications by mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2016. **113**(7): p. 1784-9.
41. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.
42. Danecek, P., et al., *Twelve years of SAMtools and BCFtools.* Gigascience, 2021. **10**(2).
43. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.
44. Ramirez, F., et al., *deepTools2: a next generation web server for deep-sequencing data analysis.* Nucleic Acids Res, 2016. **44**(W1): p. W160-5.
45. Mustachio, L.M., et al., *Targeting the SAGA and ATAC Transcriptional Coactivator Complexes in MYC-Driven Cancers.* Cancer Res, 2020. **80**(10): p. 1905-1911.
46. Numata, A., et al., *Lysine acetyltransferase Tip60 is required for hematopoietic stem cell maintenance.* Blood, 2020. **136**(15): p. 1735-1747.