

Multiple Independent Binding Sites for Small-Molecule Inhibitors on the Oncoprotein c-Myc

Dalia I. Hammoudeh,[†] Ariele Viacava Follis,[†] Edward V. Prochownik,[‡] and Steven J. Metallo^{*,†}

Department of Chemistry, Georgetown University, Washington, District of Columbia 20057, Section of Hematology/Oncology, Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania 15213, The University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania 15232, and Department of Microbiology and Molecular Genetics, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania 15213

Received January 26, 2009; E-mail: sjm24@georgetown.edu

Abstract: Deregulation of the c-Myc transcription factor is involved in many types of cancer, making this oncoprotein an attractive target for drug discovery. One approach to its inhibition has been to disrupt the dimeric complex formed between its basic helix–loop–helix leucine zipper (bHLHZip) domain and a similar domain on its dimerization partner, Max. As monomers, bHLHZip proteins are intrinsically disordered (ID). Previously we showed that two c-Myc–Max inhibitors, 10058-F4 and 10074-G5, bound to distinct ID regions of the monomeric c-Myc bHLHZip domain. Here, we use circular dichroism, fluorescence polarization, and NMR to demonstrate the presence of an additional binding site located between those for 10058-F4 and 10074-G5. All seven of the originally identified Myc inhibitors are shown to bind to one of these three discrete sites within the 85-residue bHLHZip domain of c-Myc. These binding sites are composed of short contiguous stretches of amino acids that can selectively and independently bind small molecules. Inhibitor binding induces only local conformational changes, preserves the overall disorder of c-Myc, and inhibits dimerization with Max. NMR experiments further show that binding at one site on c-Myc affects neither the affinity nor the structural changes taking place upon binding to the other sites. Rather, binding can occur simultaneously and independently on the three identified sites. Our results suggest the widespread existence of peptide regions prone to small-molecule binding within ID domains. A rational and generic approach to the inhibition of protein–protein interactions involving ID proteins may therefore be possible through the targeting of ID sequence.

Introduction

The proper biological function of most proteins requires that they interact with other proteins in complexes.^{1–3} The ability to influence the activity of a protein by interfering with such interactions by use of small organic molecules is extremely desirable, albeit challenging, since diseases often result from aberrant or failed protein–protein interactions and small molecules have significant potential as therapeutics.^{3–5} The use of low molecular weight, cell-permeable enzyme inhibitors has been very successful mainly because of the nature of the enzyme active sites, which tend to reside in well-defined cavities shielded from solvent. The dominant interactions between enzyme active sites and their specific substrates can be (and often are) mimicked by well-designed drugs.⁴ Receptor–ligand interactions are structurally similar to enzyme–substrate interactions

in that they tend to involve relatively rigid binding clefts. Enzymes and membrane receptors represent over 80% of drug targets.⁶ Only recently have other protein–protein interactions been shown to be influenced by small molecules; as recently as a decade ago, such potential was still controversial because of the typical extended surface area and flatness of protein recognition interfaces that do not have clear binding pockets.⁵

Unlike enzymes, protein–protein interaction interfaces do not provide a template for drug design and the key interacting residues are often not apparent. In at least some instances, only a small portion of the protein–protein interface contributes to high-affinity binding.⁷ This suggests that it might not be necessary for a small molecule to cover the entire surface in order to prevent a protein–protein interaction. Most known small-molecule inhibitors of protein–protein interactions for which there is a structural understanding of their binding to a target bind critical “hot spots” or functional epitopes.^{7,8}

Intrinsically disordered (ID) proteins are proteins that under physiological conditions are either completely disordered or

* Corresponding author: e-mail: sjm24@georgetown.edu.

[†] Georgetown University.

[‡] Children's Hospital of Pittsburgh, The University of Pittsburgh Cancer Institute, and University of Pittsburgh Medical Center.

(1) Gavin, A. C.; et al. *Nature* **2002**, *415*, 141–147.

(2) Ho, Y.; et al. *Nature* **2002**, *415*, 180–183.

(3) Berg, T. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 2462–2481.

(4) Hang, Y.; Hamilton, A. D. *Angew. Chem.* **2005**, *44*, 4130–4163.

(5) Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301–317.

(6) Hopkins, A. L.; Groom, C. R. *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730.

(7) Clackson, T.; Wells, J. A. *Science* **1995**, *267*, 383–386.

(8) Gardsvoll, H.; Dano, K.; Ploug, M. *J. Biol. Chem.* **1999**, *274*, 37995–38003.

contain significant regions (≥ 40 consecutive residues) of disorder.⁹ Disordered regions are characterized by extensive backbone flexibility, with transient formation of secondary structure but lacking a stable tertiary fold.¹⁰ ID proteins are prevalent in eukaryotes and are especially common in signal transduction (70% of proteins involved in signaling)⁹ and other complex regulatory pathways typical of higher organisms. Malfunction in the regulation of their activity is implicated in cancer ($\sim 80\%$ of all proteins associated with cancer)⁹ and other human diseases such as cardiovascular disease, amyloidosis, neurodegenerative disease, and diabetes.¹¹ Importantly, these proteins are biologically active in their natively disordered or ID state. The lack of a defined structure provides ID proteins certain functions or advantages that complement those of ordered proteins.^{12,13} They often participate in protein–protein or protein–nucleic acid interactions involving coupled folding and binding. These binding interactions are characterized by high specificity and modest affinity because of the entropic cost associated with their structural induction.¹⁴ ID proteins can exploit their structural flexibility to interact with different partners in one-to-many and many-to-one binding^{11,15} (thus acting as hubs) or to bind in different conformations or elicit opposing effects from binding (“moonlighting”).¹⁶ The accessible nature of ID regions makes them suitable substrates for critical posttranslational modification, and they are often involved in cell signaling pathways such as phosphorylation.^{9,14,16–18} Without the same constraints as a folded globular region, ID regions are also overrepresented as products of alternative splicing sites.¹⁹

Interest in ID proteins has been growing in recent years because of the relevant biological roles of these proteins.²⁰ The overwhelming evidence that ID proteins are functional in their disordered state has caused a reassessment of the protein structure–function paradigm.²¹ The general possibility of explicitly targeting ID protein interactions with small molecules has not been seriously considered until recently since these are not generally considered to be “druggable targets”. There are two approaches that can be considered; in the first approach, the interaction between a structured protein and an ID binding

partner is targeted. Cheng et al.²² noted that, for several of the protein–protein interactions that have been successfully modulated by small molecules, one of the partners is ID and undergoes a disorder-to-order transition upon binding to its structured partner. The authors argue that such interactions have features that allow them to be “druggable”, and since ID proteins are overrepresented in disease processes, these interactions represent a large reservoir of potential targets.²² They propose a general approach in which specific, short regions of ID sequence that are predicted to mediate protein–protein interactions (molecular recognition elements, MoREs) are used to identify their ordered binding partner, followed by structure determination of the MoRE in complex with its target and structure-based, rational drug design that replaces the MoRE with a small molecule. In the second approach, the one we deal with in this paper, a small molecule binds directly to a short segment of an ID protein, stabilizes the overall disordered state, and thereby inhibits protein–protein interactions that require coupled folding and binding (whether one or both partners are disordered).²³ Both approaches benefit energetically from avoiding the entropic penalty of folding the ID component. The second approach has a major functional advantage in that it does not require high-resolution structural data of an ID protein’s binding partner. Conversely, the second approach is not directly amenable to structure-based rational inhibitor design. However, from initial hits found via screening, structure–activity relationships can be built that lead to molecules with increased affinity as well as the development of effective pharmacophore models.^{24–26}

The *MYCC* gene product is a member of the basic helix–loop–helix leucine zipper (bHLHZip) family of transcription factors and plays important roles in cell cycle progression, cell growth, proliferation, differentiation, and apoptosis.²⁷ *MYCC* overexpression also leads to uncontrolled cell proliferation and transformation and has been shown to deregulate 10–15% of PolII-regulated genes as well as rRNA and tRNA genes regulated by RNA polymerases I and III, respectively.^{28–30} It has two independently functioning regions, the N-terminal trans-activating domain and the C-terminal bHLHZip domain.²⁷ In order to bind DNA, regulate target gene expression, and function in most biological contexts, c-Myc must dimerize with its obligate bHLHZip heterodimerization partner protein Max, which lacks a transactivation segment.^{31–33} The c-Myc–Max dimer interface is a parallel, left-handed, four-helix bundle, with each monomer composed of two α -helices separated by a loop.³³

(9) Iakoucheva, L. M.; Brown, C. J.; Lawson, J. D.; Obradovic, Z.; Dunker, A. K. *J. Mol. Biol.* **2002**, *323*, 573–584.

(10) Romero, P. R.; Obradovic, Z.; Dunker, A. K. *Appl. Bioinf.* **2004**, *3*, 105–113.

(11) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *Annu. Rev. Biophys.* **2008**, *37*, 215–46.

(12) Dunker, A. K.; Brown, C. J.; Obradovic, Z. *Adv. Protein Chem.* **2002**, *62*, 25–49.

(13) Tompa, P. *Trends Biochem. Sci.* **2002**, *27*, 527–533.

(14) Dyson, H. J.; Wright, P. E. *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60.

(15) Dunker, A. K.; Garner, E.; Guillot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J. E. *Pac. Symp. Biocomput.* **1998**, *473*, 84.

(16) Tompa, P.; Szasz, C.; Buday, L. *Trends Biochem. Sci.* **2005**, *30*, 484–489.

(17) Liu, J. G.; Perumal, N. B.; Oldfield, C. J.; Su, E. W.; Uversky, V. N.; Dunker, A. K. *Biochemistry* **2006**, *45*, 6873–6888.

(18) Dyson, J.; Martinez-Yamout, M.; Wright, P. *FEBS J.* **2005**, *272*, 360–360.

(19) Romero, P. R.; Zaidi, S.; Fang, Y.; Uversky, V. N.; Radivojac, P.; Oldfield, C. J.; Cortese, M. S.; Sickmier, M.; Le Gall, T.; Obradovic, Z.; Dunker, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8390–8395.

(20) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. *J. Mol. Recognit.* **2005**, *18*, 343–384.

(21) Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321–331.

(22) Cheng, Y.; LeGall, T.; Oldfield, C. J.; Mueller, J. P.; Van, Y.-Y. J.; Romero, P.; Cortese, M. S.; Uversky, V. N.; Dunker, A. K. *Trends Biotechnol.* **2006**, *24*, 435–442.

(23) Follis, A. V.; Hammoudeh, D. I.; Wang, H.; Prochownik, E. V.; Metallo, S. J. *Chem. Biol.* **2008**, *15*, 1149–1155.

(24) Wang, H. B.; Hammoudeh, D. I.; Follis, A. V.; Reese, B. E.; Lazo, J. S.; Metallo, S. J.; Prochownik, E. V. *Mol. Cancer Ther.* **2007**, *6*, 2399–2408.

(25) Kiessling, A.; Wiesinger, R.; Sperl, B.; Berg, T. *Chem. Med. Chem.* **2007**, *2*, 627–630.

(26) Mustata, G.; Follis, A. V.; Hammoudeh, D. I.; Metallo, S. J.; Wang, H.; Prochownik, E. V.; Lazo, J. S.; Bahar, I. *J. Med. Chem.* **2009**, *52*, 1247–1250.

(27) Dang, C. V. *Mol. Cell Biol.* **1999**, *19*, 1–11.

(28) Grandori, C.; Gomez-Roman, N.; Felton-Edkins, Z. A.; Ngouenet, C.; Galloway, D. A.; Eisenman, R. N.; White, R. J. *Nat. Cell Biol.* **2005**, *7*, 311–U121.

(29) Felton-Edkins, Z. A.; Kenneth, N. S.; Brown, T. R. P.; Daly, N. L.; Gomez-Roman, N.; Grandori, C.; Eisenman, R. N.; White, R. J. *Cell Cycle* **2003**, *2*, 181–184.

(30) Dang, C. V.; O'Donnell, K. A.; Zeller, K. I.; Nguyen, T.; Osthus, R. C.; Li, F. *Semin. Cancer Biol.* **2006**, *16*, 253–264.

The c-Myc monomer is ID and forms a stable secondary and tertiary structure only upon the coupled binding and folding transition in association with Max.

The c-Myc protein is a relevant yet challenging target for drug discovery.^{34–37} It is overexpressed in the majority of human cancers, and in the vast majority of cases it harbors no mutations to distinguish it from c-Myc expressed by nontransformed cells.^{27,38} Several small-molecule inhibitors of the c-Myc–Max interaction have been reported in the literature over the past 6 years.^{39–45} In a recent review, Berg⁴⁶ summarized the current knowledge of c-Myc–Max dimerization inhibitors and those of other transcription factors. Using a high-throughput assay, our group (Yin et al.⁴⁵) identified seven structurally unrelated yet highly specific small-molecule inhibitors of the c-Myc/Max interaction, (Figure 1) from a 10 000 member library where the molecules were selected to cover the broadest part of the biologically relevant pharmacophore space. These inhibitors were conspicuous for their compliance to Lipinski's rules.⁴⁷

Using a fluorescence polarization (FP) assay, we previously showed that two c-Myc–Max inhibitors with fluorescent properties identified by Yin et al.⁴⁵ (10058-F4 and 10074-G5), as well as improved derivatives of 10058-F4, bound directly to the ID c-Myc monomer, which remained disordered post-binding.^{23,24} We deciphered the binding sites of 10058-F4 and 10074-G5 by monitoring their association with truncations and point mutations of the c-Myc bHLHZip domain. As a final means of verification, 10058-F4 was found to bind to a synthetic peptide composed of residues c-Myc_{402–412}, while 10074-G5 bound to the c-Myc_{363–381} segment.²³ The residues of these peptides directly involved in compound binding were then identified from NMR studies of the complexes. These binding sites, residues 402–409 for 10058-F4 and 366–375 for 10074-G5, are composed of only a small number of amino acids each (~10); the local nature of these binding interactions might therefore reduce the entropic cost associated with structural induction upon binding. Moreover, the binding of these two compounds to their corresponding sites can happen simultaneously and independently. Here we show that these two distinct binding sites also interact with four of the five other low molecular weight inhibitors first identified by Yin et al.⁴⁵ despite the chemical

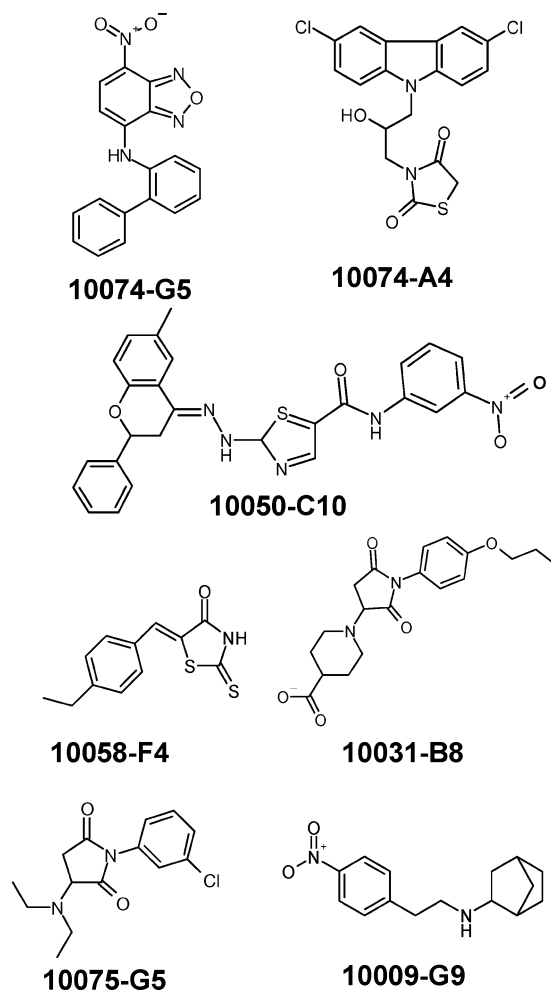


Figure 1. Structures of the seven c-Myc–Max small-molecule inhibitors found by Yin et al.⁴⁵

diversity of these compounds. In addition we identify a third, independent binding site on the c-Myc bHLHZip domain for the fifth compound, 10074-A4. Binding to the three sites on the ID monomer can occur simultaneously and independently. Moreover, binding to one site does not interfere with binding to other sites.

Materials and Methods

Recombinant Protein Expression and Purification. Truncations of c-Myc bHLHZip were overexpressed in *Escherichia coli* BL21DE3(pLysS). The DNA sequence that codes for c-Myc_{353–437} was cloned in to the expression vector pET151D-TOPO. This vector codes for a hexahistidine (6× His) tag, which is separated by a tobacco etch virus (TEV) protease digestion site from the N-terminus of the insert. The DNA sequence of the c-Myc gene was modified from the c-Myc/pET SKB3 construct kindly supplied by Dr. S. K. Nair (University of Illinois, Urbana–Champaign). The c-Myc_{353–405}/pET SKB3 construct was obtained through insertion of a stop codon by use of QuickChange (Stratagene). The other constructs, c-Myc_{370–409}, c-Myc_{380–439}, cMyc_{390–439}, and c-Myc_{400–439}, were generated via polymerase chain reaction (PCR) amplification excluding different portions of the gene, followed by cloning into pET151D-TOPO vector (Invitrogen).²³

The 6× His-tagged human Max(p21) and Max(p22), the 151 and 160 amino acid isoforms of Max, were also cloned into

- (31) Blackwood, E. M.; Eisenman, R. N. *Science* **1991**, *251*, 1211–1217.
 (32) Prendergast, G. C.; Lawe, D.; Ziff, E. B. *Cell* **1991**, *65*, 395–407.
 (33) Nair, S. K.; Burley, S. K. *Cell* **2003**, *112*, 193–205.
 (34) Hermeking, H. *Curr. Cancer Drug Targets* **2003**, *3*, 163–175.
 (35) Prochownik, E. V. *Expert Rev. Anticancer Ther.* **2004**, *4*, 289–302.
 (36) Darnell, J. E., Jr. *Nat. Rev. Cancer* **2002**, *2*, 740–749.
 (37) Gibbs, J. B. *Science* **2000**, *287*, 1969–1973.
 (38) Nesbit, C. E.; Tersak, J. M.; Prochownik, E. V. *Oncogene* **1999**, *18*, 3004–3016.
 (39) Kiessling, A.; Sperl, B.; Hollis, A.; Eick, D.; Berg, T. *Chem. Biol.* **2006**, *13*, 745–751.
 (40) Mo, H.; Henriksson, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6344–6349.
 (41) Xu, Y.; Shi, J.; Yamamoto, N.; Mossa, J. A.; Vogt, P. K.; Janda, K. D. *Bioorg. Med. Chem.* **2006**, *14*, 2660–2673.
 (42) Berg, T.; Cohen, S. B.; Desharnais, J.; Sonderegger, C.; Maslyar, D. J.; Goldberg, J.; Boger, D. L.; Vogt, P. K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3830–3835.
 (43) Bagnasco, L.; Tortolina, L.; Biasotti, B.; Castagnino, N.; Ponassi, R.; Tomati, V.; Nieddu, E.; Stier, G.; Malacarne, D.; Parodi, S. *FASEB J.* **2007**, *21*, 1256–1263.
 (44) Pescarolo, M. P.; Bagnasco, L.; Malacarne, D.; Melchiori, A.; Valente, P.; Millo, E.; Bruno, S.; Basso, S.; Parodi, S. *FASEB J.* **2001**, *15*, 31–33.
 (45) Yin, X.; Giap, C.; Lazo, J. S.; Prochownik, E. V. *Oncogene* **2003**, *22*, 6151–9.
 (46) Berg, T. *Curr. Opin. Chem. Biol.* **2008**, *12*, 1–8.
 (47) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.

the pET151D-TOPO vector (Invitrogen), and overexpressed in *E. coli* strain BL21(DE3*). Bacterial cultures of all c-Myc and Max constructs were grown at 37 °C in Luria–Bertani (LB) medium to $OD_{600} \approx 0.8$, then induced with 0.5 mM isopropyl thio- β -D-galactoside (IPTG) for 5 h. Cultures were harvested and lysed in a buffer containing 8 M urea, 100 mM NaH_2PO_4 , and 10 mM Tris; pH 8.0. Proteins were purified on a nickel–nitrilotriacetic acid (Ni-NTA) agarose column (Qiagen) with pH gradient elution and then were desalted. The 6 \times His tag of each expressed protein was cleaved by use of TEV protease [previously expressed in a pET24 vector (from S. K. Nair) and purified on NTA-Ni–agarose under native conditions]. All c-Myc-derived proteins and both Max isoforms were further purified by reverse-phase HPLC (Vydac C18) using a water/acetonitrile gradient containing 0.1% trifluoroacetic acid (TFA) and then lyophilized. Protein concentrations were determined by measurement of OD_{280} .

Synthesis and Purification of c-Myc_{402–412} and c-Myc_{363–381} Peptides. The c-Myc_{402–412} and c-Myc_{363–381} peptides were prepared as described previously.²³ Briefly, the c-Myc_{402–412} peptide was synthesized at the University of Delaware in the laboratory of Professor Neal Zondlo via standard 9-fluorenylmethoxycarbonyl (Fmoc) solid-phase synthesis on a Rainin PS3 automated peptide synthesizer. The peptide was synthesized as the C-terminal amide (Rink amide resin) and was acetylated at the N-terminus while still attached to the resin. The peptide was cleaved from the resin by use of 92.5% TFA, 2.5% triisopropylsilane, and 5% water. The c-Myc_{363–381} peptide was synthesized by the University of Vermont Protein Core Facility and was delivered as the acetylated and amidated lyophilized crude peptide. Each peptide was redissolved in water, filtered, and purified to homogeneity by reverse-phase HPLC (Vydac C18) with a water/acetonitrile gradient containing 0.1% TFA. The concentration of c-Myc_{402–412} was determined by its absorbance at 280 nm. The concentration of c-Myc_{363–381}, which lacks appreciable absorbance at 280 nm, was assessed by integration of the HPLC 215 nm peak against a bovine serum albumin standard after normalization of the extinction coefficients of the two peptides.

Synthesis of 10058-F4 and 10074-G5. 10058-F4 [(5Z)-5-[(4-ethylphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one]⁴⁸ and 10074-G5 [7-nitro-N-(2-phenylphenyl)-2,1,3-benzoxadiazol-4-amine]⁴⁹ were synthesized and purified according to established procedures and were characterized by ¹H and ¹³C NMR on a 300 MHz Varian INOVA spectrometer. The other small-molecule inhibitors, 10074-A4, 10009-G9, 10031-B8, 10075-G5, and 10050-C10, were purchased from ChemBridge Inc. (San Diego, CA.) and their purity was assessed by HPLC to be 90%, 97%, 91%, 95%, and 94% respectively.

Fluorescence Polarization Competition and Fluorescence Measurements. Samples were prepared containing 10 μ M equimolar mixtures of either 10058-F4 or 10074-G5 and c-Myc_{353–437} in the presence of an excess concentration (40 or 100 μ M) of one of the nonfluorescent compounds (10009-G9, 10031-B8, 10074-A4, or 10075-G5) in 1 \times PBS, pH 7.4, 1 mM dithiothreitol (DTT), and 5% dimethyl sulfoxide (DMSO). The compound 10050-C10, due to its poor solubility, was tested at 20 μ M in the presence of 20% DMSO. The samples were analyzed in a Photon Technology International QuantaMaster

fluorometer (Birmingham, NJ) equipped with polymer sheet polarizers at an excitation wavelength of 380 nm and an emission wavelength of 468 nm for samples containing 10058-F4 or at 470 and 550 nm, respectively, for samples containing 10074-G5. Triplicate experiments were performed at 25 °C with sample-specific G-factor determination and background correction. Competition affinity experiments were performed over a range of concentrations (3–200 μ M) of the nonfluorescent inhibitor being titrated against either 10058-F4 or 10074-G5 in the presence of c-Myc_{353–437}. Reported data represent the average of 3–5 independent experiments. Data from the competition experiments were fit as described previously²³ by use of the following equation, which requires an equimolar ratio between c-Myc_{353–437} and 10058-F4 or 10074-G5:

$$\text{pol} = \text{pol}_0 + \Delta\text{pol} \left(\frac{-\frac{[I]}{[M]} - 1 + \sqrt{\left(\frac{[I]}{[M]} + 1\right)^2 + 4\frac{[I]}{[M]}(K_{\text{comp}} - 1)}}{2K_{\text{comp}} - 2} \right) \quad (1)$$

where pol_0 is the polarization in the absence of competing inhibitor, Δpol is the total change in polarization after complete disruption of the complex between c-Myc and fluorescent inhibitor, $[I]$ is the concentration of nonfluorescent inhibitor (variable), $[M]$ is the concentration of c-Myc_{353–437} and 10058-F4 or 10074-G5, and K_{comp} represents the dissociation constant between c-Myc and the nonfluorescent inhibitor over the dissociation constant between c-Myc and the fluorescent inhibitor.

The affinity between 10074-G5 and c-Myc_{353–437} in the presence of 100 μ M 10074-A4 was determined by monitoring the fluorescence polarization of 10074-G5 upon serial dilution of an equimolar mixture of c-Myc_{353–437} and 10074-G5 in the presence of constant 10074-A4. Three independent samples in 1 \times PBS were analyzed at 25 °C and data were fit to the following equation:

$$\frac{[\text{complex}]}{[C]_0} = \frac{2 + \frac{K_{\text{obs}}}{[C]_0} - \sqrt{\left(-2 - \frac{K_{\text{obs}}}{[C]_0}\right)^2 - 4}}{2} \quad (2)$$

The fraction of complex was fitted to the polarization end points of the titration, $[C]_0$ represents the concentration of 10074-G5 and c-Myc_{353–437}, and K_{obs} represents the fitted dissociation constant parameter.^{23,50}

Circular Dichroism. Samples of c-Myc_{402–412} or c-Myc_{363–381} (20 μ M) in the absence and presence of an equal or excess concentration of inhibitor (20 μ M 10058-F4, 50 μ M 10031-B8, 50 μ M 10075-G5, or 100 μ M 10009-G9 binding to c-Myc_{402–412} and 20 μ M 10050-C10 binding to c-Myc_{363–381}) were prepared in 1 \times PBS (pH 7.4). The inhibitors were added from 10 mM stock solutions in ethanol. Spectra were recorded at 25 °C in a 1 mm path-length quartz cuvette on a Jasco J710 spectropolarimeter and were averaged from five independent samples. Similar conditions and concentrations were employed to monitor the effect of compound 10074-A4 on the spectra of several truncated versions of c-Myc bHLHZip (Myc residues 353–437, 353–405, 370–409, 380–439, 390–439, and 400–439). A titration of c-Myc_{370–409} in the presence of the inhibitor was performed by serial dilution of 1:1 mixtures of 10074-A4 and c-Myc_{370–409} with monitoring at 245 nm. The experimental data were fit to eq 2, where the polarization end

(48) Khodair, A. J. *Heterocycl. Chem.* **2002**, *39*, 1153–1160.

(49) Summers, W. A.; Lee, J. Y.; Burr, J. G. *J. Org. Chem.* **1975**, *40*, 1559–1560.

(50) Park, S. H.; Raines, R. T. *Methods Mol. Biol.* **2004**, *261*, 161–166.

points had been replaced by ellipticity end points. Samples for competition experiments of 10074-A4 against Myc–Max(p21) heterodimers were prepared in 1× PBS (pH 7.4). Solutions of c-Myc_{353–437} were incubated for 1 h with varying concentrations of inhibitor; 1.5 μM Max(p21) was then added and reactions were further incubated for 20 min. Samples for the competition experiment of 10074-A4 (100 μM) against the Max(p22) homodimer (1.5 μM) were prepared in 1× PBS (pH 7.4). Triplicate samples were analyzed after a 20 min incubation time. Readings for the competition experiment were performed at 25 °C in a 1 cm path-length cuvette by monitoring ellipticity at 222 nm. Data were converted to MRW and plotted versus inhibitor concentration. Experimental data were fitted to the form of eq 1, where the polarization end points had been replaced by ellipticity end points.

NMR Spectroscopy. Experiments were performed on a 500 MHz Varian INOVA instrument equipped with a 5 mm single nucleus indirect detection probe. All experiments were performed at 25 °C. Samples containing c-Myc_{370–409} in the absence and presence of 10074-A4 or c-Myc_{353–437} in the absence and presence of 10058-F4 or 10074-G5 only; 10058-F4 and 10074-G5; 10074-G5 and 10074-A4; or 10058-F4 and 10074-G5 and 10074A4 were prepared in 100% D₂O and 5 mM sodium phosphate buffer, pH 7.5. Samples of the pure proteins were also prepared in 90% H₂O–10% D₂O and 5 mM sodium phosphate buffer, pH 6.3. All samples used for nuclear Overhauser enhancement spectroscopy (NOESY) experiments were degassed by sonication and purged with nitrogen. The final peptide concentration was ~0.5 mM. Each small molecule was added to peptide solutions from a 0.1 M stock in DMSO-*d*₆ to a final concentration of ~1 mM. Complete backbone ¹H resonance assignments for c-Myc_{370–409} and partial assignments for c-Myc_{353–437}²³ were obtained from total correlation spectroscopy (TOCSY) and H_α(*i*)–H_N(*i* + 1) NOEs of low pH, 90% H₂O–10% D₂O samples. ¹³C resonance assignments for c-Myc_{370–409} were obtained from heteronuclear multiple quantum coherence (HMQC) of 100% D₂O samples, pH 7.5. Spectra of the pure peptides were not affected beyond the disappearance of amide and amine signals within the explored pH range. All spectra were acquired over similar sweep widths of ~10 ppm in *t*₁ and *t*₂ for homonuclear experiments or ~140 ppm in *t*₁ for HMQC experiments, and sizes of 512–1024 × 2048 complex points. Correlation spectroscopy (COSY) spectra were acquired with water suppression by use of selective presaturation and 16–32 scans per *t*₁ increment and a relaxation delay of 1.5–2.5 s. TOCSY spectra of the pure peptides were acquired, with water suppression by use of TNOCSY on-resonance presaturation, 64 scans per *t*₁ increment, a relaxation delay of 1.5–2.5 s and a TOCSY mixing time of 60 ms. NOESY spectra were acquired with TNNOESY presaturation, 64 scans per *t*₁ increment, a relaxation delay of 1.5 or 2.5 s, and a NOESY mixing time of 150 ms. HMQC spectra were acquired on natural ¹³C abundance samples via on-resonance water presaturation, 128 scans per *t*₁ increment, and a relaxation delay of 2.0 s, with a C–H one-bond coupling constant of 140 Hz. Spectra were processed with MestReC software (Mestrelab Research, Santiago de Compostela, Spain). Data were filled by linear prediction to a final Fourier transform size of 2048 × 2048 points and weighted by sine square and sine bell apodization over *t*₁ and *t*₂, respectively, before Fourier transformation.

Molecular Modeling of c-Myc_{370–409} Segment in Free and Bound States. Molecular models based on NMR ¹H and ¹³C chemical shift information were generated as reported previ-

ously.²³ Approximate Φ–Ψ backbone and χ₁ side-chain angles of the free peptides and of the peptide in complex with 10074-A4 were obtained from ¹Hα, ¹³Cα, and ¹³Cβ (and ¹HN for the free peptides) chemical shift values by use of the web server PREDITOR.⁵¹ Predicted dihedral angles and confidence levels are reported in Table S1 in Supporting Information. The dihedral predictions, in combination with scwrl3⁵² software for assessment of side-chain conformation, were employed to generate free and bound input structures, which were minimized for 10 000 time steps in an automatically generated cubic water box (PSFgen) by use of CHARMM27⁵³ parameters implemented in NAMD2⁵⁴ software. The minimized conformer structures displayed no bad parameters after a PROCHECK validation test.⁵⁵ Docking between the complex conformer structure of the peptide and both enantiomers of 10074-A4 was performed with the AutoDock LGA algorithm.⁵⁶ A test docking of 25 runs was performed with a 50-point side cubic energy grid with 1 Å/point resolution to assess pose clustering. A 60-point side cubic energy grid with 0.375 Å/point resolution, centered on the expected binding site (based on NMR information), was then used for energy scoring in the final docking. A total of 10 docking runs with an initial population of 150 random conformations were performed with 2 500 000 energy evaluations each. Selected side-chain rotamers, chosen upon experimental indications and results of preliminary rigid docking, and all the rotating bonds of 10074-A4 were kept flexible during docking. The following side chains were unconstrained except for χ₁ rotation: Phe₃₇₄, Leu₃₇₇, Arg₃₇₈, Gln₃₈₀, Ile₃₈₁, Glu₃₈₃, Leu₃₈₄, and Glu₃₈₅. The best poses were geometrically optimized by use of UFF parameters⁵⁶ to 0.1 kcal mol⁻¹ Å⁻¹ convergence. The final complex models were validated with PROCHECK analysis.

Results

Binding Site and Affinity Determination. To characterize the binding of the c-Myc inhibitors 10031-B8, 10075-G5, 10009-G9, 10050-C10, and 10074-A4,⁴⁵ knowledge of the previously determined binding sites of the two fluorescent compounds 10058-F4 and 10074-G5, which bind to two independent segments of the bHLHZip domain (c-Myc_{353–437}), was exploited. The exact protein segments involved in the binding of 10058-F4 and 10074-G5 had been obtained from mutagenesis and truncation studies on c-Myc_{353–439}.²³ When bound to c-Myc_{353–437}, the fluorescence polarization of 10058-F4 or 10074-G5 is higher than that of the free compound. We used a competition assay in which an excess concentration of each of the nonfluorescent compounds with unknown binding sites was added to a solution of c-Myc_{353–437} containing an equimolar concentration of either 10058-F4 or 10074-G5. A decrease of the fluorescence polarization signal of 10058-F4 from ~0.12 to ~0.01 upon addition of excess competing compound showed that compounds 10031-B8, 10075-G5, and 10009-G9 displaced

(51) Berjanskii, M. V.; Neal, S.; Wishart, D. S. *Nucleic Acids Res.* **2006**, *34*, W63–W69.

(52) Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L., Jr. *Protein Sci.* **2003**, *12*, 2001–2014.

(53) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(54) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(55) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

(56) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

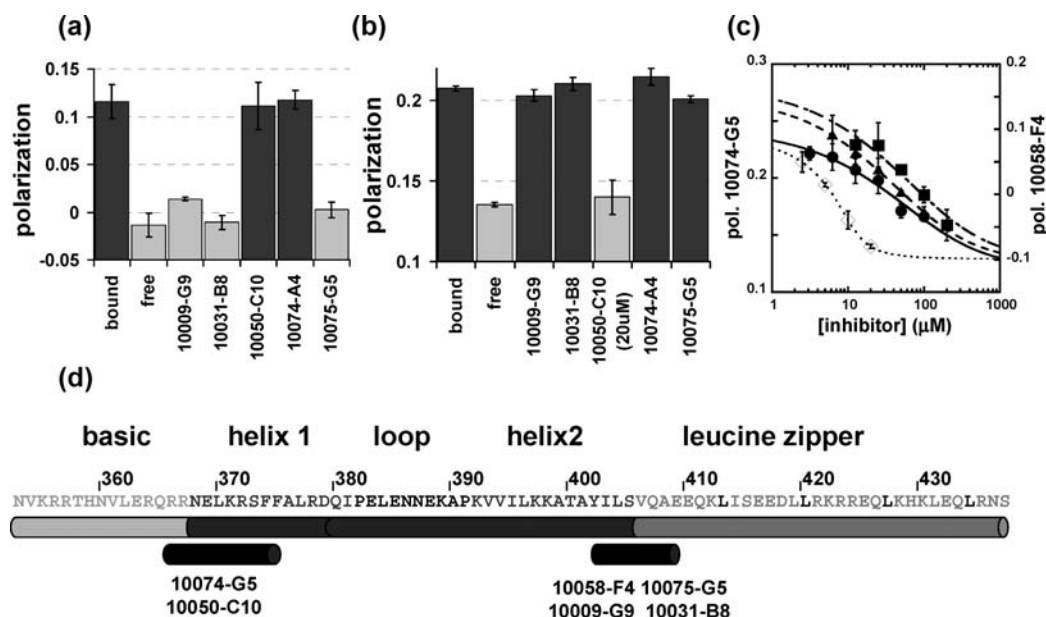


Figure 2. Fluorescence polarization (FP) competition assays: (a) FP of 10058-F4 either bound to c-Myc_{353–437} (bound), alone (free), or in the presence of both peptide and various compounds as indicated. (b) FP of 10074-G5 either bound to c-Myc_{353–437} (bound), alone (free), or in the presence of both peptide and various compounds as indicated. (c) FP competition titration of the nonfluorescent 10009-G9 (■), 10031-B8 (●), or 10075-G5 (▲) at varying concentrations (3.1–200 μM) against 10058-F4/c-Myc_{353–437} (right-hand axis) and of 10050-C10 (◇) (3.1–20 μM) against 10074-G5/c-Myc_{353–437} (left-hand axis). (d) Binding sites of small-molecule inhibitors including the two fluorescent index compounds 10058-F4 and 10074-G5.

Table 1. Tabulated Affinities (K_d or K_{obs}) of Seven Small-Molecule Inhibitors

compound	Myc _{353–437} affinity (μM)	K_{obs} (μM)	
		from 10058-F4 competition	from 10074-G5 competition
10058-F4	5.3 ± 0.7 ^a	NB	NB
10074-G5	2.8 ± 0.7 ^a	NB	NB
10009-G9		40 ± 10	NB
10075-G5		24 ± 4	NB
10031-B8		16 ± 4	NB
10050-C10		NB	0.9 ± 0.3
10074-A4	21 ± 2 ^b	NB	NB

^a By direct binding to c-Myc FP titration from ref 36. ^b c-Myc_{370–409} affinity from ICD titration. NB, no binding.

10058-F4 from its binding site on c-Myc_{353–437} (Figure 2a). Similarly, 10074-G5 was displaced from its binding site on c-Myc by 10050-C10 (Figure 2b). Neither 10031-B8, 10075-G5, nor 10009-G9 displaced 10074-G5 from c-Myc, nor did 10050-C10 displace 10058-F4. That 10074-A4 did not displace either of the two fluorescent compounds suggested that it bound a third, as yet unidentified site on the bHLHZip sequence of c-Myc.

To determine the affinity of the inhibitors for their target sequences on c-Myc, competition titrations using fluorescence polarization were performed and the resulting data for the titrated inhibitor and a constant concentration of 10058-F4 or 10074-G5 was fit to a competition constant parameter (K_{comp}), which corresponds to the ratio between the dissociation constant (K_d) between c-Myc and the titrated inhibitor and that between c-Myc and 10058-F4 or 10074-G5 (Figure 2c). The K_d values between c-Myc and 10031-B8, 10075-G5, 10009-G9, and 10050-C10 were calculated (Table 1) from the previously determined c-Myc affinities of 10058-F4 and 10074-G5.²³ Compounds that bound to the 10058-F4 site did so with 3–8-fold lower affinity than 10058-F4, while 10050-C10 (the largest compound) bound with 3-fold higher affinity than 10074-G5 to its binding site (Table 1).

Changes of Secondary Structures of Binding Sites upon Small-Molecule Binding.

Only small changes were induced in the CD spectra of c-Myc_{353–437} upon complex formation with 10058-F4 and 10074-G5.²³ This is most likely due to the fact that conformational changes upon small-molecule binding involve only a small number of residues within the much larger bHLHZip domain. The CD spectra indicated that c-Myc_{353–437} retained its predominantly ID structure after complex formation.²³ Synthetic peptides encompassing the two known binding sites, c-Myc_{402–412} (the site for 10058-F4 binding) and c-Myc_{363–381} (the site for 10074-G5 binding),²³ were therefore used for subsequent CD studies. A change in the spectrum of c-Myc_{402–412} was observed upon addition of 10031-B8, 10075-G5, and 10009-G9. The conformational changes induced on this peptide by these three compounds were most noticeable in the reduced intensity or disappearance of a strong negative peak around 207 nm, indicative of a random coil or dynamic conformation. The overall effect on the CD spectrum of c-Myc_{402–412} was consistent upon addition of the three inhibitors (Figure 3a) and is similar to what we previously observed with 10058-F4.²³ The compound 10009-G9 resulted in a less pronounced change compared to the other compounds. The spectrum of Myc_{363–381} was also significantly altered by the binding of 10050-C10, most noticeably by the disappearance of the strong negative peak at 207 nm (Figure 3b), which is also similar to what was observed upon binding of 10074-G5 to Myc_{363–381}.²³ These results confirm the binding-site determination from the FP data and show the highly localized nature of these binding interactions. All the compounds discussed above alter the structure of the peptide to which they bind, although the extent and type of change varies.

Binding Site of 10074-A4. The nonfluorescent compound 10074-A4 was unable to displace either 10058-F4 or 10074-G5 from their respective binding sites, thus excluding them as the sites for interaction with this compound. We confirmed the activity of 10074-A4 by competition against Max for c-Myc

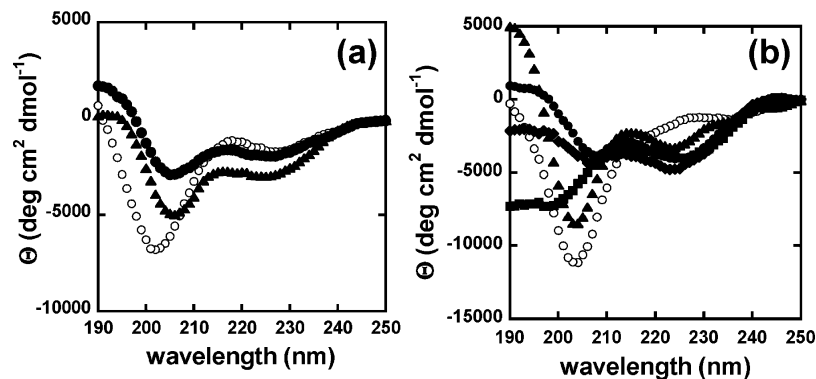


Figure 3. Circular dichroism (CD) spectra of c-Myc_{363–381} and c-Myc_{402–412} upon addition of inhibitors. (a) CD spectra of c-Myc_{363–381} alone (○), with 10074-G5 (●), and with 10050-C10 (▲). (b) CD spectra of c-Myc_{402–412} alone (○), with 10058-F4 (●), with 10009-G9 (▲), with 10075-G5 (■), and with 10031-B8 (◆).

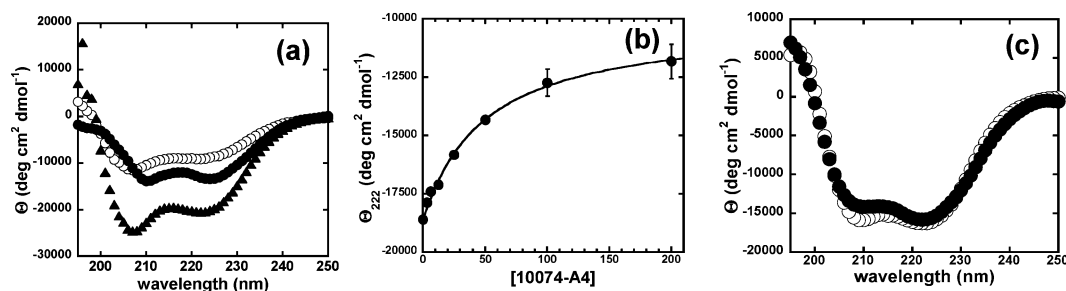


Figure 4. Disruption of the c-Myc–Max dimer upon addition of 10074-A4: (a) Summation of the two CD spectra: c-Myc and Max(p21) taken independently (○) and CD spectra of the c-Myc–Max(p21) dimer before the addition of 10074-A4 (▲) and after the addition of 10074-A4 (●). (b) Competition titration by CD of the c-Myc–Max dimer upon addition of 10074-A4. Error bars represent SEM. (c) CD spectra of Max(p22) in the absence (○) and presence (●) of 10074-A4.

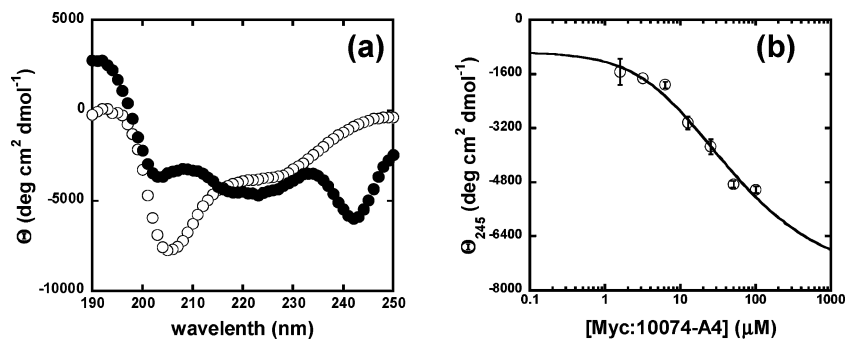


Figure 5. CD spectra of c-Myc_{370–409} upon addition of 10074-A4, showing the induced CD band with a minimum at 245 nm. (a) CD spectra of c-Myc_{370–409} (○) and c-Myc_{370–409} after addition of 10074-A4 (●). (b) Titration of the ICD band at 245 nm by serial dilution of a 1:1 mixture of 10074-A4 and c-Myc_{370–409}. Error bars represent SEM.

binding in a CD experiment where the extent of c-Myc–Max dimer formation was monitored by measuring the ellipticity at 222 nm, indicative of α -helical content (Figure 4b). The Max(p21) isoform (which homodimerizes only at very high micromolar concentrations)⁵⁷ was used in this experiment in order to avoid convoluting the effect on α -helical content by formation of Max homodimer. The compound was able to displace Max with an observed K_{comp} of $32 \pm 3 \mu\text{M}$ which, based on the independently determined c-Myc–Max affinity, corresponds to an inhibitor affinity of $13 \pm 3 \mu\text{M}$. A similar experiment performed with Max(p22), an isoform with high affinity for homodimer formation, showed that 10074-A4 was unable to disrupt the Max homodimer, thus confirming its

specific binding to c-Myc (Figure 4c). Information about the location of the binding site for 10074-A4 was obtained by monitoring the CD of various truncated c-Myc bHLHZip peptides upon addition of this compound. We monitored the effect of 10074-A4 on the CD spectra of peptides encompassing the truncated segments systematically: c-Myc_{353–405} (Figure S2C in Supporting Information), c-Myc_{370–409} (Figure 5a), c-Myc_{380–439} (Figure S2D in Supporting Information), c-Myc_{390–439} (Figure S2E in Supporting Information), c-Myc_{400–439} (Figure S2F in Supporting Information), and c-Myc_{363–381} (Figure S2G in Supporting Information). The spectra of c-Myc_{370–409} and c-Myc_{353–405} show a significant change after the addition of 10074-A4, whereas the CD spectrum of c-Myc_{353–437} shows a very limited change, retaining its ID structure after the addition of 10074-A4 (Figure S2B in Supporting Information). The spectra are indicative of a binding interaction involving only a

(57) Zhang, H.; Fan, S. J.; Prochownik, E. V. *J. Biol. Chem.* **1997**, *272*, 17416–17424.

short segment of the protein. The change in the spectra of the peptides versus the spectrum of the entire bHLHZip domain is once again consistent with a simple averaging effect, that is, a higher relative effect of local structure rearrangements due to complex formation on a peptide composed of fewer amino acids than the full c-Myc bHLHZip domain. An induced circular dichroism (ICD) effect⁵⁸ on the small-molecule absorption band centered at 245 nm was also observed in these two spectra. ICD is a phenomenon observed in interactions between a chiral molecule (a peptide in this case) and an achiral or racemic (as in this case) compound, where the chiral surroundings affect the absorption transition of the compound. Alternatively, enantiomer-specific effects on the racemic compound's extinction coefficient or wavelength shifts, or both, effectively lead to the enantiomers' optical resolution as a consequence of their diastereoselective interaction with the chiral component. Little change and no ICD band were observed in the spectrum of the c-Myc_{380–439} peptide, and no changes in the spectra of c-Myc_{390–439}, c-Myc_{400–439}, and c-Myc_{363–381} were observed. These results clearly show that the peptide with the sequence more narrowly spanning the binding site is c-Myc_{370–405}. Monitoring the intensity of the ICD band upon serial dilution of a 1:1 mixture of 10074-A4 and c-Myc_{370–409}, we defined a binding curve and determine the complex affinity as $21 \pm 2 \mu\text{M}$ (Figure 5b). The direct affinity measurements and the affinity determined by Myc–Max disruption are in reasonable agreement with each other.

NMR Studies of 10074-A4 Binding. In order to more specifically characterize the structural features of the binding interaction between 10074-A4 and its deduced binding site, NMR spectroscopy was employed on c-Myc_{370–409}. The backbone ¹H assignments for the pure peptide were obtained from $\alpha\text{H}_i\text{--NH}_{(i+1)}$ NOESY cross peaks; proton information was then mapped onto a ¹H–¹³C HMQC spectrum to obtain ¹³C α assignments. Addition of an excess racemic mixture of 10074-A4 to the peptide (both present at concentrations >10-fold above their dissociation constants) caused changes in backbone chemical shifts of residues predominantly in the helix-1 and loop regions, suggesting that the exact location of the interaction site for this compound is near but C-terminal to that of 10074-G5 (Figure 8a and Figure S3 in Supporting Information). The increased cross peaks in the NOESY spectra of the bound peptide compared to the free peptide indicate some degree of structural induction upon complex formation with 10074-A4 (Figure 6a). The free peptide showed very limited cross peaks; upon addition of compound, additional interresidue cross peaks are present, half of which are three residues away (Figure 6b). Intermolecular cross peaks in the NOESY spectrum of the complex confirmed the presence of hydrophobic interactions between the inhibitor and hydrophobic groups on the peptide (residues Leu377, Ile381, and Leu384; interactions are also observed with Arg378 and Asp 379 residues) located in this region (Figure 6b and Figure S4 in Supporting Information). Two patterns of intrapeptide cross peaks between residues located three positions away from each other in the sequence suggested some extent of helical conformation in the segments spanning Phe374–Gln380 and Lys392–Leu396 (Figure 6b). However, the overall number of interresidue and intermolecular NOE cross peaks was limited even in the bound state (24 in total, six of which involved residues directly adjacent to each other), indicating a high extent of residual flexibility.

Backbone chemical shift information was used to assess secondary structure trends in the peptide by means of chemical shift indexing (CSI)⁵⁹ (Figure 8 and Figure S3 in Supporting Information). Chemical shift information from ¹H α , ¹³C α , and ¹³C β (and ¹HN for the free peptides) was employed to generate dihedral constraints for molecular modeling with the PREDITOR program.⁵¹ The paucity of structurally relevant NOE signals (and their complete absence in the free peptide) meant actual structure determination was not possible. Inclusion of the limited NOEs present would unduly bias and distort the resulting models and were used instead to corroborate chemical shift data.^{60–62} In both the free and the bound state, the peptide is highly flexible and secondary structure elements may be transient and can be observed only locally as indicated by NOE, chemical shift, and CD data. By use of the backbone angles from PREDITOR, models of a likely average conformation of the peptide in its bound state and more dynamic free state were generated. These models were expected to provide an assessment of the overall topology of the peptide in the two conformations; however, they are not meant to, and cannot, define their detailed structural features. The ¹H and ¹³C CSI of the free peptide reveal a pattern of mixed downfield (typical of β -sheet structures) and upfield (typical of α -helices) shifts with respect to random coil values, alternating with segments of residual helical content, as also indicated by the ¹³C CSI. Such a pattern, considered typical of coil conformations, could be associated with regions displaying residual structure in the presence of local conformational constraints, as opposed to a more dynamic random coil state, where the backbone chemical shifts would more consistently match the expected random coil values.⁶³ The models of the peptide in its free and bound states generated from dihedral constraints suggest the formation of a cavity at the N-terminus of the loop region, flanked by Phe374/375, Ala376, and Leu377 in a helical conformation at the N-terminus of the helix-1 sequence (Figure 6c,d). Although determined completely independently, the model of the bound state is highly consistent with the indication of α -helical segments from NOE cross peaks. Comparison of the free and bound models indicates that the relative repositioning of two segments, roughly corresponding to residues from the helix-1 and loop regions, generates a conformation favorable to binding. Molecular docking of the inhibitor to the bound model suggests a possible mode of binding of the compound to the described site favored mainly by a series of hydrophobic interactions. There are an unusually high number of hydrophobic residues in the H1-loop segment of c-Myc bHLHZip compared to the entire domain (7 out of 11 amino acids, or 64% in this segment, versus 31 out of 85, or 36% in the entire domain). The docking of both enantiomers displayed a similar mode of binding and similar docking scores (0.3 kcal mol⁻¹ in favor of the *S* enantiomer). Poses for both enantiomers are overall consistent with the independently generated NOE data. This simulation provided a general understanding of the binding interaction but cannot generate precise binding information or identification of a favored binding enantiomer.

(58) Allenmark, S. *Chirality* **2003**, *15*, 409–422.

(59) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647–1651.

(60) Williamson, M. P.; Waltho, J. P. *Chem. Soc. Rev.* **1992**, *21*, 227–236.

(61) Wishart, D. S.; Sykes, B. D. *Methods Enzymol.* **1994**, *239*, 363–392.

(62) Eliezer, D. *Curr. Opin. Struct. Biol.* **2009**, *19*, 23–30.

(63) Rose, G. D.; Fleming, P. J.; Banavar, J. R.; Maritan, A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16623–16633.

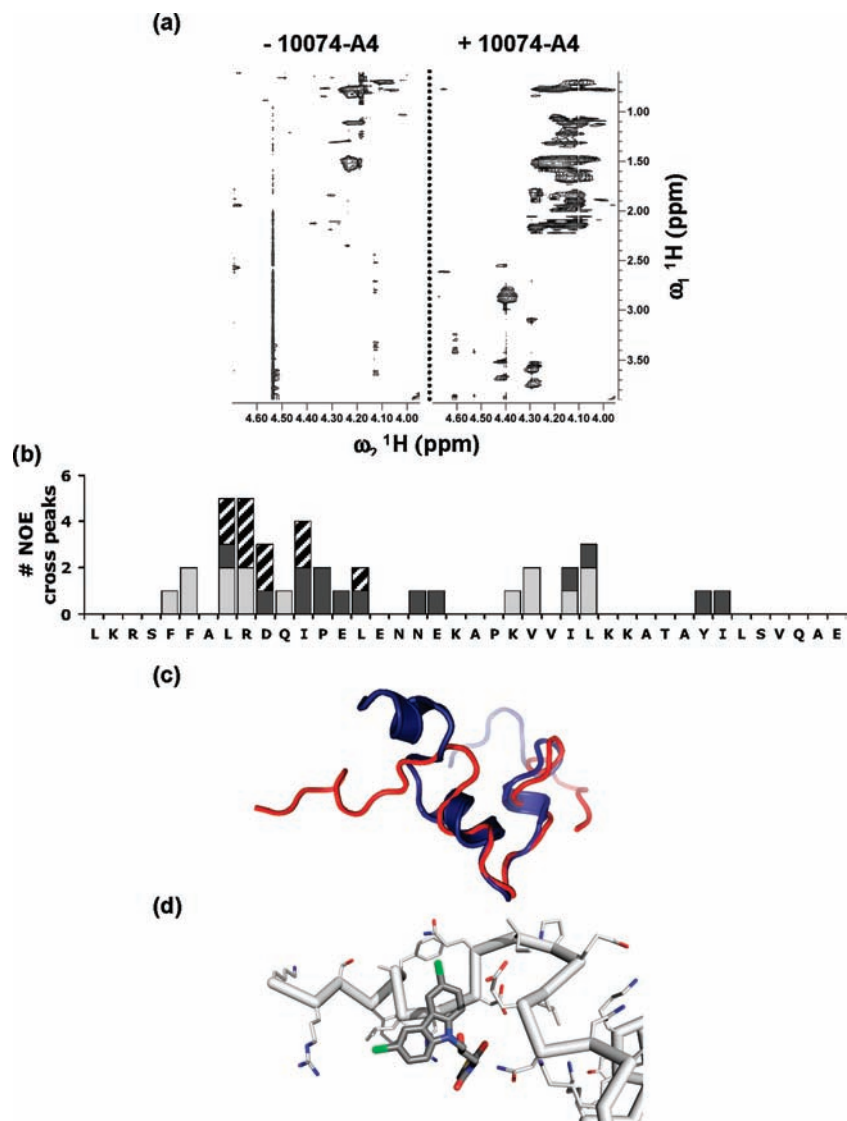


Figure 6. (a) NOESY spectra showing (left) the $^1\text{H}\alpha$ region of c-Myc $_{370-409}$ and (right) the $^1\text{H}\alpha$ region of c-Myc $_{370-409}$ after the addition of 10074-A4. (b) Sequence mapping of interresidue and intermolecular NOESY cross peaks of the complex: between one and two residues away (dark gray), three residues away (light gray), and intermolecular cross peaks (hatched). (c) Models of the free (blue) and bound (red) peptide to 10074-A4. (d) Docking of 10074-A4 on c-Myc $_{370-409}$.

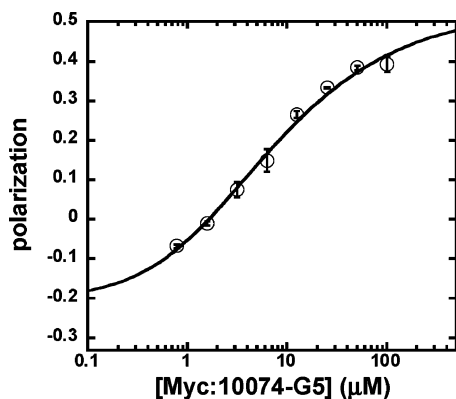


Figure 7. Equimolar binding titration of 10074-G5 and c-Myc $_{353-437}$ ($10\ \mu\text{M}$) in the presence of $100\ \mu\text{M}$ 10074-A4. Error bars represent SEM.

Simultaneous Independent Binding of Multiple Sites on c-Myc. The NMR study of 10074-A4 binding indicated that its interaction site on c-Myc was adjacent to that of the compound

10074-G5. We therefore tested whether the binding of one of these compounds to c-Myc would alter the affinity of the other. Binding between the fluorescent compound 10074-G5 and the c-Myc bHLHZip domain was monitored by FP upon serial dilution of a 1:1 protein inhibitor mixture in the presence of a constant excess ($100\ \mu\text{M}$) of the nonfluorescent compound 10074-A4 (Figure 7). The affinity value observed ($K_{\text{obs}} = 3 \pm 1\ \mu\text{M}$) for 10074-G5 under these conditions was the same as its c-Myc affinity in isolation ($K_d = 2.8 \pm 0.7\ \mu\text{M}$), demonstrating that the binding of this compound was not affected by the presence of 10074-A4. This result indicated that the binding sites on c-Myc are truly independent and that even two adjacent and closely spaced binding sites could independently interact with their target inhibitor, most likely because of the scarcity of medium- or long-distance conformational constraints in the disordered state of the protein. This conclusion was supported by NMR studies on c-Myc $_{353-437}$ in its free form and upon the sequential addition of the three inhibitors 10058-F4, 10074-G5, and 10074-A4. Samples were also analyzed that contained

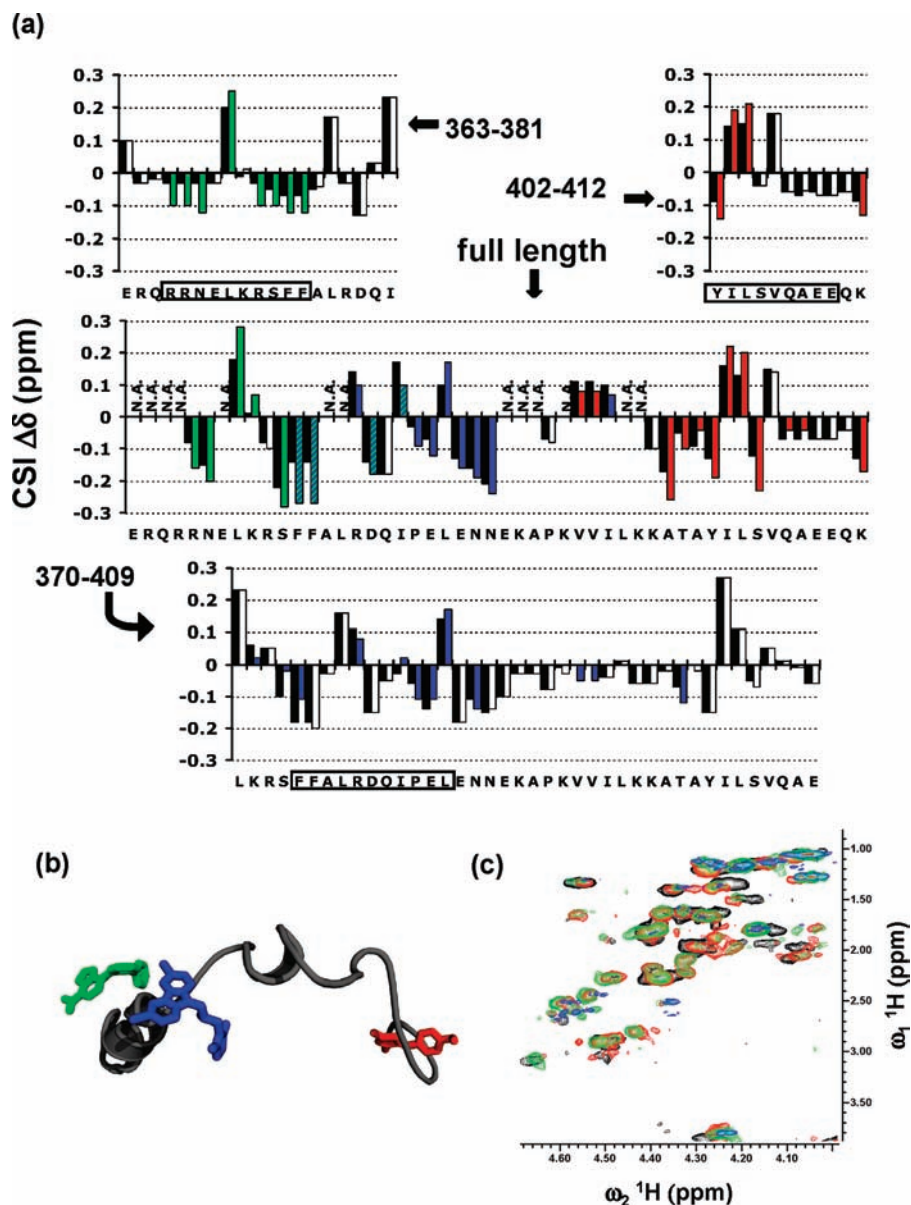


Figure 8. Simultaneous binding on bHLHZip c-Myc_{353–437}. (a, top left) ¹H CSI plot of the free (black) and bound (green) c-Myc_{363–381} upon addition of 10074-G5 with the binding site marked on the sequence. (a, top right) ¹H CSI plot of free (black) and bound (red) of c-Myc_{402–412} upon addition of 10058-F4 with binding site marked on the sequence. (a, bottom) ¹H CSI plot of the c-Myc_{370–409} in its free (black) and bound to 10074-A4 (blue) states. In the center of the figure is the ¹H CSI plot of c-Myc_{353–437} only (black) and upon the addition of 10058-F4, 10074-G5, and 10074-A4. Only residues that display a change greater than 0.02 in chemical shift upon binding are displayed as colored bars in the CSI plots of the bound peptides. (b) Model of the simultaneous binding of 10074-F4 (red), 10074-G5 (green), and 10074-A4 (blue). (c) Overlay of COSY spectra of c-Myc_{353–437} (black) upon addition of 10074-F4 (red), upon addition of 10074-F4 and 10074-G5 (green), and upon addition of 10074-F4, 10074-G5, and 10074-A4 (blue).

c-Myc_{353–437} and 10074-G5 with further addition of 10074-A4. Partial ¹H backbone assignments for c-Myc_{353–437} were obtained, taking advantage of the low redundancy of some amino acids found within or in proximity to each binding site (Ser₃₇₃, Phe₃₇₄, Phe₃₇₅, Asp₃₇₉, Pro₃₈₂, Thr₄₀₀, Tyr₄₀₂, and Ser₄₀₅). Chemical shift changes similar to those observed upon complex formation with peptide segments spanning the isolated binding sites (c-Myc_{402–412} for 10058-F4, c-Myc_{363–381} for 10074-G5, and c-Myc_{370–409} for 10074-A4) were observed in the full c-Myc bHLHZip domain upon binding of each compound. The changes in chemical shift induced by each binding event appeared to be independent from each other, and only residues Phe₃₇₄, Phe₃₇₅, Asp₃₇₉, and Ile₃₈₁ were affected by both 10074-G5 and 10074-A4 binding (Figure 8) although, as determined by affinity

measurements, there is no competition for binding between the two compounds (Figure 7). These experiments provided partial sequence mapping of the conformational rearrangements caused by each compound's binding to the full-length c-Myc bHLHZip and support the idea of local, independent binding interactions.

Discussion

Within the disordered 85 amino acid bHLHZip domain of c-Myc we have identified three distinct binding sites that recognize seven structurally diverse small-molecule inhibitors. The first of these sites, composed of amino acids 402–409 (YILSVQAE), bound four of the seven originally identified c-Myc–Max selective inhibitors. This sequence on c-Myc is disordered in the monomer but would be situated at the interface

between the H2 and Zip region in the c-Myc–Max dimer (Figure 2d). A previous analysis of the c-Myc bHLHZip disorder propensity²³ performed with the VSL2 algorithm⁶⁴ from the PONDR set, indicated that this segment lies at the interface between a small region of reduced disorder probability and a more extended region of predicted disorder. This site is able to bind to different, structurally unrelated compounds using the same sequence of amino acids and seems to have high plasticity (e.g., the ability to bind multiple, chemically distinct ligands), most likely as a consequence of its lack of stable folding or secondary structure. By analogy, proteins involved in signaling, regulation, and transcription use a single short ID sequence to bind to different partners. These proteins can also use multiple disordered regions to flexibly bind several structured proteins at the same time.^{9,14,16,65}

We additionally found that compound 10019-D3 (Figure S1A in Supporting Information), a so-called “dual specific” inhibitor of both c-Myc–Max and Id2-E47 (Yin et al.)⁴⁵ binds to the same c-Myc_{400–409} site as 10058-F4 (Figure S1C in Supporting Information). 10019-D3 has the same pyrazolo[1,5-*a*]pyrimidine-2-carboxamide core structure as Mycro1, -2, and -3, which were compounds found by Berg et al.^{39,46} to inhibit the c-Myc–Max interaction.^{25,39} Because 10019-D3 is fluorescent, we were able to determine its binding to c-Myc_{353–437} through a FP assay and to calculate a K_d of $11 \pm 4 \mu\text{M}$ (Figure S1B in Supporting Information). Since it is the pyrazolo[1,5-*a*]pyrimidine-2-carboxamide core that seems to bind to c-Myc, it is likely that Mycro1, -2, -3, and structurally related pyrazolo[1,5-*a*]pyrimidines bind to c-Myc_{402–412}. These results suggest that various independent screens for c-Myc–Max disruptors are likely to identify active compounds that also bind to one of the three identified sites on Myc. Given the diversity of structures found to bind the 402–409 site from a relatively small 10 000 compound library,⁴⁵ it is reasonable that other diversity libraries would also contain scaffolds capable of binding this site. Recently, a three-dimensional pharmacophore model of c-Myc compounds binding at this site was developed and demonstrated to be capable of predicting additional inhibitors with diverse structures that bound with affinities similar to 10058-F4.²⁶ ID sequences that have affinity for small-molecule binding show a dramatic ability to recognize many different chemical structures, which should facilitate finding small-molecule binders. However, the specificity of the binders must be confirmed by counterscreens against nontargeted proteins, as shown by Berg and co-workers³⁹ and Yin et al.⁴⁵

The three remaining compounds, 10074-G5, 10050-C10, and 10074-A4, bind to adjacent sites located predominantly in the H1 region of the c-Myc bHLHZip domain. A sharp reduction in disorder propensity was observed, by use of the VSL2 algorithm,⁶⁴ in the portion of c-Myc sequence encompassing both binding sites.²³ The second site on c-Myc_{353–437} encompassing amino acids 366–375 (Figure 2d) bound 10050-C10, which is structurally different from 10074-G5. This further supports the plasticity of ID binding sites. A third site on cMyc_{353–437}, centered on residues 375–385, bound 10074-A4. This interaction was confirmed by means of CD and NMR spectroscopy. The backbone chemical shift pattern of the binding site of the peptide is overall conserved after complex formation, suggesting that the binding event

may be favored by the presence of segments with residual structure within the disordered protein. Even though the second and third sites are adjacent to each other on the c-Myc primary sequence and overlap at two residues (F375 and F376), the different compounds bind simultaneously and independently to their target sequence without any evident interference (Figure 8).

The three binding sites comprise short sequences of amino acids on an ID protein that are recognized by and bind to small organic molecules. Only a few examples of small molecules specifically recognizing short disordered sequences of amino acids on proteins have been reported. A derivative of Taxol was shown to bind to short disordered peptides similar in sequence to the disordered loop region on Bcl-2, thus leading the authors to identify and confirm Bcl-2 as a Taxol-binding protein.⁶⁶ Recently, while investigating the action of γ -secretase modulators, Kukar et al.⁶⁷ found small molecules that selectively bind to a short stretch of amino acids ($A\beta$ 28–36) on the amyloid precursor protein. Morohashi et al.⁶⁸ found that a known antitumor drug (NK109) bound to the same short sequence of amino acids (PNXXXXP) on multiple protein targets. These results and examples clearly show that short ID sequences can be targeted by small organic druglike molecules. The current understanding of druggable targets excludes ID sequences.^{69,70} However, as we continue to see examples of small-molecule binding to ID sequences and begin to understand the affinity and specificity of these interactions, ID regions should be considered as potential druggable targets.

Upon interaction with a rigid protein binding site, changes in the structure of the small molecule will have large effects on the binding affinity, and only a narrow region of chemical space constitutes the best match ligand for any given binding pocket.⁷¹ This concept underlies the ability to optimize hits by structure–activity relationship (SAR) analysis, fragment-based approaches, and various docking techniques. When binding to ID domains, structurally very different molecules can bind to the same sequence on the protein while retaining specificity for it. This likely explains why many of the 10058-F4 derivatives previously reported by us, and synthesized without knowledge of their binding site on c-Myc, were as active or more active than the parental compound.

ID domains are known to bind to their structured protein partners with high specificity and low affinity.¹⁴ Similar criteria, involving enthalpy–entropy tradeoffs, as well as structural plasticity, seem to be implicated in their binding to small molecules. Within the still relatively narrow literature regarding small-molecule inhibition of protein–protein interactions, even fewer examples of ID protein targets, such as c-Myc, have been reported. Studies that report small-molecule binding to ID protein targets were designed to test for a detectable end effect of the inhibition and could not provide a structural and mechanistic understanding of it.³⁹ Their outcome indicates that when screening for the inhibition of ID proteins, a blind screen may actually be targeting multiple different segments of a target protein. The present results suggest that a rational approach to

(64) Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Dunker, A. K. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 176–182.
(65) Oldfield, C. J.; Meng, J.; Yang, J. Y.; Yang, M. Q.; Uversky, V. N.; Dunker, A. K. *BMC Genomics* **2008**, *9*, S1.

(66) Rodi, D. J.; Janes, R. W.; Sanganeer, H. J.; Holton, R. A.; Wallace, B. A.; Makowski, L. *J. Mol. Biol.* **1999**, *285*, 197–203.

(67) Kukar, T. L.; et al. *Nature* **2008**, *453*, 925–929.

(68) Morohashi, K.; Yoshino, A.; Yoshimori, A.; Saito, S.; Tanuma, S.; Sakaguchi, K.; Sugawara, F. *Biochem. Pharmacol.* **2005**, *70*, 37–46.

(69) Imming, P.; Sinning, C.; Meyer, A. *Nat. Rev. Drug Discovery* **2006**, *5*, 821–834.

(70) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. *Nat. Rev. Drug Discovery* **2006**, *5*, 993–996.

the inhibition of ID protein–protein interactions may be possible through an appropriate analysis of a target protein sequence. We found that small-molecule binding sites in ID proteins have certain sequence criteria. They can be found in predicted regions of low disorder, contain nonconserved residues, and tend to have higher hydrophobic content than the rest of the sequence.²³ Small molecules capable of modulating ID protein function might be found by screening libraries of compounds for binding to small segments of the target protein selected for their sequence characteristics. This technique has the advantage of inherently defining the binding site along the protein sequence. The structural plasticity of complexes between ID proteins and small molecules demonstrated here suggests that this approach might be applied broadly and result in reasonable hit rates.

Acknowledgment. This work was sponsored by NIH Grant R01-CA105033 to E.V.P. and a Young Investigator Award from the American Cancer Society to S.J.M.

Supporting Information Available: Figures showing the structure of 10019-D3, FP titration of 10019-D3/c-Myc_{353–437}, and FP competition of 10019-D3 against 10058-F4; CD spectra of 10074-A4 and various truncations of c-Myc; overlay of TOCSY and HMQC spectra of c-Myc_{370–409} free and bound to 10074-A4 in the $\alpha^1\text{H}$ –side chain and $\alpha^1\text{H}$ – $\alpha^{13}\text{C}$ regions; $\alpha^{13}\text{C}$ and $\beta^{13}\text{C}$ chemical shift indexing of c-Myc_{370–409} free and bound to 10074-A4; tables containing NMR assignments and prediction output from PREDITOR; and complete references 1, 2, 53, and 57. This information is available free of charge via the Internet at <http://pubs.acs.org>.

JA900616B

(71) Xu, H. E.; Lambert, M. H.; Montana, V. G.; Plunket, K. D.; Moore, L. B.; Collins, J. B.; Oplinger, J. A.; Kliewer, S. A.; Gampe, R. T.; McKee, D. D.; Moore, J. T.; Willson, T. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13919–13924.