Specific inhibition of formation of transcription complexes by a calicheamicin oligosaccharide: A paradigm for the development of transcriptional antagonists

STEFFAN N. Ho*, SERGE H. BOYER†, STUART L. SCHREIBER‡, SAMUEL J. DANISHESKY†§, AND GERALD R. CRADBREE¶

*Department of Pathology and Howard Hughes Medical Institute, Beckman Center for Molecular and Genetic Medicine, Stanford University School of Medicine, Stanford, CA 94305; †Department of Chemistry, Yale University, New Haven, CT 06511; ‡Department of Chemistry, Harvard University, Oxford Street, Cambridge, MA 02138; and §Laboratory of Bio-organic Chemistry, Sloan Kettering Institute for Cancer Research, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021

Contributed by Samuel J. Danishefsky, May 4, 1994

ABSTRACT  Sequence-specific DNA ligands that antagonize DNA–protein interactions represent a potentially powerful means of modulating gene expression. Calicheamicin γ1, a member of the DNA-cleaving enediyne class of anticancer antibiotics, binds to specific DNA sequences through an aryltetrascaccharide domain. To take advantage of this unique sequence-specific recognition capability, the methyl glycoside of the aryltetrasaccharide of calicheamicin γ1 (CLM-MG) was used to investigate the ability of glycoconjugate DNA ligands to inhibit DNA–protein interactions. CLM-MG inhibits the formation of DNA–protein complexes at micromolar concentrations in a sequence-specific manner and rapidly dissociates preformed complexes. CLM-MG also inhibits transcription in vivo with similar sequence specificity. These results suggest a strategy for the development of a class of novel biological probes and therapeutic agents.

The ordered regulation of gene expression underlying specific developmental programs or cellular responses to physiologic stimuli relies on sequence-specific DNA–protein interactions (1). This requisite DNA sequence specificity suggests a means to control cellular responses through the use of DNA ligands that interfere with these specific DNA–protein interactions. To be biologically useful such ligands would need to (i) bind DNA with a high degree of sequence specificity, (ii) antagonize DNA–protein interactions either by inhibition of DNA–protein complex formation or by displacement of preformed complexes, and (iii) exhibit sufficient hydrophobicity to pass through cell membranes and thereby function in vivo. Several types of DNA ligands have been shown to satisfy some of these criteria, including oligodeoxynucleotides involved in triple-helix formation (2, 3), peptide nucleic acids (4), and nonintercalating minor groove DNA ligands such as netropsin and distamycin A (5). However, none of these adequately satisfy all the requirements for specific and biologically useful transcriptional antagonists.

The enediyne compound calicheamicin γ1 (CLM; Fig. 1A), representative of a class of enediyne glycoconjugate DNA ligands, binds to and cleaves DNA at specific 4-bp sequences. Double-strand DNA cleavage occurs through bioreduction of the trisulfide and subsequent 1,4-benzenoid diradical formation, which results in hydrogen abstraction from adjacent nucleotide bases (10–13). Affinity cleavage experiments (7, 8) demonstrate that CLM preferentially cleaves DNA at cytosine-containing homopyrimidine tracts, preferentially TTCT. Comparisons of the DNA cleavage specificity of CLM derivatives indicate that the aryltetrascaccharide domain is largely responsible for sequence-specific DNA binding (8, 14–18). Furthermore, CLM appears to bind in the minor groove of DNA, on the basis of the observation that DNA scission occurs at sites on each strand that are separated by 3 bp (ref. 18; Fig. 1B).

To further investigate sequence-specific carbohydrate ligands as possible transcriptional antagonists, we have synthesized the methyl glycoside of the aryltetrasaccharide moiety of CLM (CLM-MG), using glycol assembly methodology (ref. 6; Fig. 1A). This compound lacks the enediyne component and therefore does not cleave DNA. Previous DNA footprinting analysis indicated that CLM-MG exhibits DNA sequence-specific binding that is similar to that of the parent compound (17). Here we demonstrate that CLM-MG not only inhibits sequence-specific DNA–protein interactions in vitro but also rapidly displaces preformed DNA–protein complexes. Furthermore, CLM-MG inhibits transcription in vivo with DNA sequence specificity, thereby providing a paradigm for the development of glycoconjugate DNA ligands as biologically specific transcriptional antagonists.

MATERIALS AND METHODS

Preparation of Nuclear Extracts and DNA-Binding Assays. Nuclear extracts were prepared as previously described (19) from TAg Jurkat cells (20) stimulated for 2 hr with 2 μM ionomycin plus phorbol 12-tetradecanoate 13-acetate (PTA) at 20 ng/ml, HeLa cells stimulated for 2 hr with PTA at 20 ng/ml, or MDCK cells. CLM-MG was solubilized in 10% (vol/vol) EtOH or 2% (vol/vol) dimethyl sulfoxide (DMSO) to a maximum concentration of 10 mM (for reference, this resulted in an extinction coefficient of 24,600 ± 1300 M⁻¹cm⁻¹ at a λmax of 213 nm in aqueous solution containing <0.01% DMSO). Electrophoretic mobility-shift assays (EMSAs) were performed by incubating various concentrations of CLM-MG or the appropriate buffer control with 0.1 ng of DNA probe (=20,000 cpm; see below) for 15 min at room temperature, followed by the addition of 2 μl (5–10 μg) of nuclear extract, in a final volume of 15 μl. The incubation was continued for 45 min at room temperature, and samples were subsequently loaded onto nondenaturing gels. The binding buffer consisted of 10 mM Tris-HCl (pH 7.6), 80 mM NaCl, 1 mM EDTA, 0.5 mM dithiothreitol, 5% (vol/vol) glycerol, and 2.5 μg of poly (dI-dC). Protein–DNA complex formation was quantitated using an AMBIS radioanalytic imaging system. The following DNA sequences were utilized as probes for the gel-shift assays: NFAT, 5′-GATCTAGGAGGAAAACGTGTTTCAGT-3′; CLM, calicheamicin γ1; CLM-MG, methyl glycoside of the aryltetrasaccharide moiety of CLM; PTA, phorbol 12-tetradecanoate 13-acetate; IL2, interleukin 2.

Abbreviations: CLM, calicheamicin γ1; CLM-MG, methyl glycoside of the aryltetrasaccharide moiety of CLM; PTA, phorbol 12-tetradecanoate 13-acetate; IL2, interleukin 2.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

GATC; API, 5'-TCGAGTCACGCAGTCG; NF-xB, 5'-AGGGATTTCAC; Oct-1, 5'-TGTAAATATGTAAAA-CATTCTG; Oct-1/OAP, 5'-TCTGAAATAAGTGTAAATATGTAAAA-CATTCTG; NF-1, 5'-CAAACTGTCAATAATATATTAAGGAG; and Sp1, 5'-TCGAGGGCGGGTCTGA.

**Reportor Gene Assays.** Tag Jurkat cells (10^7) were transiently transfected by electroporation (Bio-Rad Gene Pulser; 960 μF, 250 V, 0.4-cm cuvette width) with 4 μg of reporter plasmids containing the human interleukin 2 (IL2) minimal promoter with various upstream enhancer sequences directing the expression of secreted alkaline phosphatase (21, 22). Cells were harvested after approximately 24 hr and aliquoted into 96-well flat-bottom microtiter plates (2 × 10^4 cells per well in 100 μl of complete medium). Various concentrations of CLM-MG or the appropriate control buffer were added to duplicate wells in 5 or 10 μl. After 60 min at 37°C, iomycin and PTA were added in 100 μl of complete medium to final concentrations of 2 μM and 20 ng/ml, respectively. Indicated concentrations of CLM-MG represent the final concentration in the 200-μl culture.

**RESULTS**

**CLM-MG Inhibits Formation of DNA-Protein Complexes.** The effect of CLM-MG on the binding of the transcription factors NFAT (23), API (24), Oct-1 (25) and Oct-1-associated protein (OAP) (26), NFκB (27), HNF-1 (28), and Sp1 (29) was measured by electrophoretic mobility-shift assay. These protein complexes are representative members of several distinct classes of transcription factors, including the basic-leucine zipper or bZIP class (API), the rel/dorsal class (NFκB), the homeodomain class that contains a helix-turn-helix motif (HNF1, Oct-1), and the zinc finger class (Sp1).

Predicted CLM-MG binding sites are present within four of the six DNA sequences to which these transcription factors bind (Fig. 2A). Among these DNA sequences, the presumably preferred CLM-MG binding site TCTT is present only in the NFAT recognition sequence.

The addition of CLM-MG to DNA-protein binding reactions resulted in the complete inhibition of NFAT-DNA complex formation at concentrations >30 μM CLM-MG, with half-maximal inhibition occurring at approximately 5 μM CLM-MG. In contrast, there was no significant inhibition of API or Sp1 complex formation at concentrations of CLM-MG that significantly inhibited NFAT-DNA complex formation and minimal inhibition at concentrations as high as 100 μM (Fig. 2B and 2C). The formation of DNA-protein complexes involving NFκB, Oct-1, OAP, and HNF-1 was also inhibited by CLM-MG, although to a much lesser extent than formation of the NFAT complex (Fig. 2B). These results indicate that CLM-MG is capable of inhibiting DNA-protein complex formation in vitro and are consistent with the results predicted on the basis of presence or absence of CLM-binding sites within the transcription factor recognition sequences (Fig. 2A).

**CLM-MG Rapidly Dissociates Preformed DNA-Protein Complexes.** CLM-MG rapidly dissociated preformed NFAT-DNA complexes, with >80% inhibition of NFAT-DNA complex formation observed within 1 min of the addition of 100 μM CLM-MG (Fig. 3). In contrast, addition of excess unlabelled DNA probe resulted in no significant dissociation of the DNA-protein complex within 1 min (Fig. 3A, lane 7). This rapid dissociation of preformed DNA-protein complexes by CLM-MG indicates that sensitivity to inhibition by CLM-MG is not dependent on differences in DNA-protein complex dissociation rates. In addition, the similarity between the rate of dissociation of preformed NFAT-DNA induced by CLM-MG at room temperature and the apparent rate at 4°C...
(Fig. 3B) is consistent with the rate constant for the association of CLM-MG and DNA being diffusion limited.

**CLM-MG Inhibits Formation of DNA–Protein Complexes with DNA Sequence Specificity.** To provide further evidence that the inhibition of DNA–protein complex formation by CLM-MG is sequence specific, alterations were made in the putative CLM-MG-binding sites within the NFAT recognition sequence (Fig. 4A). Dissociation of preformed DNA–protein complexes was measured after incubation with CLM-MG for 1 min to minimize differences in the dissociation rates of NFAT with the altered sequences. Substitution of thymidine for guanosine in the first guanosine dinucleotide resulted in decreased sensitivity of the NFAT complex to inhibition by CLM-MG (Fig. 4B). Alterations in the second guanosine dinucleotide eliminated NFAT binding (data not shown), consistent with the identification of these nucleotides as contact sites by methylation interference studies (J. P. Shaw and G. R. C., unpublished data). The presence of overlapping CLM-MG-binding sites within the NFAT recognition sequence gives rise to the prediction that increasing the distance between these sites may increase sensitivity to CLM-MG, since overlapping sites may result in decreased occupation of the site more directly involved in protein contacts. Comparison between the wild-type NFAT recognition sequence and an altered sequence containing an inserted adenosine nucleotide shows that binding of NFAT to the latter sequence is more sensitive to inhibition by CLM-MG (Fig. 4). Addition of excess unlabeled DNA probe for 1 minute (Fig. 4A, lane 5) resulted in less than 10% dissociation of the DNA–protein complex for each of the DNA sequences, indicating that the differential inhibition by CLM-MG is not simply a function of differences in the rate of dissociation of NFAT for the different DNA-binding sequence. Thus, the extent to which NFAT–DNA complex formation is inhibited by CLM-MG is dependent on the presence of CLM-MG-binding sites, providing further evidence that CLM-MG functions to inhibit DNA–protein interaction by binding to DNA in a sequence-specific manner.

**CLM-MG Inhibits Transcription in Vivo with DNA Sequence Specificity.** Finally, to determine whether the sequence-specific inhibition of DNA–protein interactions by CLM-MG observed in vitro could be manifest in vivo, the effect of CLM-MG on the expression of the secreted alkaline phosphatase reporter gene directed by multimerized NFAT, AP1, NFκB, or Oct1/OAP recognition sequences located upstream of the IL2 minimal promoter was measured in transient transfection assays. Expression of each of these reporter constructs requires stimulation of the transfected cells by cross-linking the T-cell antigen receptor or by treatment with ionomycin and/or PTA. CLM-MG markedly inhibited the induced expression of the reporter gene directed by the NFAT-IL2 promoter with a half-maximal inhibitory concentration of approximately 60 μM (Fig. 5). In contrast, expression of the reporter gene directed by the API-IL2 promoter showed no inhibition by CLM-MG up to a concentration of 125 μM, but rather was increased by CLM-MG, indicating that the inhibition of NFAT-directed reporter gene expression is not due to nonspecific cell toxicity. The augmentation in API-directed gene expression may be due to
were subject to substitution 1), CLM-MG DNA-protein data. The probe NFAT of was directed reporter gene to resulting conditions was at bition over manner with less than significantly and HNF1-ML2 promoters recognition site, AP1 taining of AP1 from release of at room experiments, with error bars indicating the range of the values. For dissociation of DNA–protein complexes at 4°C, complexes were formed at room temperature for 60 min and transferred to 4°C for 15 min prior to the addition of CLM-MG; the gels were run at 4°C. Binding in the presence of 100-fold excess unlabeled NFAT probe for 1 min was not significantly different from control.

release of AP1 from composite DNA-binding complexes containing AP1 that bind to sequences containing a CLM-MG recognition site, such as the NFAT complex. The NFκB-IL2 and HNF1-IL2 promoters were also inhibited, although significantly less than inhibition of the NFAT-IL2 promoter (Fig. 5).

CLM-MG inhibited cell proliferation in a dose-dependent manner with approximately 75% inhibition of [3H]thymidine incorporation over a 24-hr period at 250 μM and 50% inhibition at 125 μM (data not shown). Cell viability under these conditions was approximately 80% and 100%, respectively, according to cell counts and trypan blue dye exclusion. At concentrations resulting in half-maximal inhibition of NFAT-directed reporter gene expression (approximately 60 μM), there was no effect on cell viability and approximately 25% inhibition of [3H]thymidine incorporation over a 24-hr period.

The antiproliferative and toxic effects of CLM-MG are not unexpected and likely result from specific DNA binding and concomitant inhibition of gene expression, given that the sequence specificity is limited to 4-bp sequences.

DISCUSSION

Transcription factors typically form interactions with DNA along the major groove (1). CLM, in contrast, appears to occupy the minor groove (19). Assuming the DNA–protein interactions that are sensitive to inhibition by CLM-MG depend primarily on major groove contacts, the rapid dissociation of preformed DNA–protein complexes by CLM-MG (Fig. 3) is unlikely to result from direct competition for a shared binding site. Rather, CLM-MG binding in the minor

Fig. 3. Rapid dissociation of preformed NFAT–DNA complexes by CLM-MG. (A) Representative gel shift. DNA–protein complexes were allowed to form on 60 min at room temperature. Samples were removed before and 15 s, 30 s, 1 min, 2 min, and 4 min after addition of CLM-MG (100 μM) (lanes 1–6, respectively) and immediately loaded onto a nondenaturing polyacrylamide electrophoresis gel. Lane 7 represents a sample loaded 1 min after the addition of 100-fold excess unlabeled NFAT probe. fp, Free probe. (B) Quantitation of gel-shift data for dissociation of DNA–protein complexes by CLM-MG performed at room temperature (●) or 4°C (○). Each point represents the mean of two separate experiments, with error bars indicating the range of the values. For dissociation of DNA–protein complexes at 4°C, complexes were formed at room temperature for 60 min and transferred to 4°C for 15 min prior to the addition of CLM-MG; the gels were run at 4°C. Binding in the presence of 100-fold excess unlabeled NFAT probe for 1 min was not significantly different from control.

The antiproliferative and toxic effects of CLM-MG are not unexpected and likely result from specific DNA binding and concomitant inhibition of gene expression, given that the sequence specificity is limited to 4-bp sequences.

DISCUSSION

Transcription factors typically form interactions with DNA along the major groove (1). CLM, in contrast, appears to occupy the minor groove (19). Assuming the DNA–protein interactions that are sensitive to inhibition by CLM-MG depend primarily on major groove contacts, the rapid dissociation of preformed DNA–protein complexes by CLM-MG (Fig. 3) is unlikely to result from direct competition for a shared binding site. Rather, CLM-MG binding in the minor

Fig. 4. Inhibition of NFAT–DNA complex formation by CLM-MG is determined by CLM-MG binding sequences. (A) Representative gel-shift data. The NFAT DNA probe sequences used are shown (in part) to the left of each gel-shift panel. The top sequence represents wild-type NFAT probe sequence. The middle and bottom sequences contain sequence changes (indicated in bold) that alter putative CLM-MG-binding sites. DNA–protein complexes were allowed to form at room temperature for 60 min and transferred to 4°C. After 15 min, control buffer (lane 1), CLM-MG at 10^{-4}, 10^{-4.5}, or 10^{-5} M (lanes 2–4, respectively), or 100-fold excess unlabeled probe (lane 5) was added. After 1 min the samples were subject to electrophoresis. (B) Quantitation of gel-shift data. ●, Wild-type NFAT probe sequence (n = 3); ■, A insertion (n = 4); ○, G-to-T substitution (n = 4). Points represent mean; error bars indicate SEM.
groove may either induce a DNA conformation that is incompatible with protein binding along the major groove or render the specific region of DNA more rigid, thereby prohibiting the DNA conformational changes or bending that have been demonstrated to occur upon protein binding (30) and that may be required for functional protein–DNA complex formation. This proposed mechanism by which CLM inhibits DNA–protein interactions is supported by the recent structural characterization of a CLM-DNA complex by NMR (31), in which Walker et al. found that DNA conformation is distorted upon binding CLM, while the conformation of CLM itself appears to exhibit minimal alterations.

This hypothetical binding interaction contrasts with that of the minor groove ligand netropsin, which binds triple-helix DNA without displacing the major groove-bound third strand (32), or 1-methylimidazole-2-carboxamide netropsin, which can bind simultaneously with the major groove-binding protein GCN4(226-281) at a common binding site (33). However, distamycin, an amide pyrrole minor groove ligand like netropsin, is capable of inhibiting DNA–protein interactions (34, 35) and dissociating preformed DNA–protein complexes (36). Furthermore, evidence for DNA conformational changes induced by distamycin is provided by direct structural studies of distamycin–DNA complexes (37, 38). Thus, targeting the minor groove, which appears to be accessible to small ligands in the presence of protein interacting along the major groove, may be a useful strategy in the design of DNA-binding ligands.

The preferential inhibition of NFAT-directed reporter gene expression (Fig. 5) reflects in vivo the same specificity of inhibition of DNA–protein complex formation by CLM-MG that was observed in vitro (Fig. 2) and is consistent with CLM-MG inhibiting gene expression in vitro by inhibiting protein–DNA complex formation in a sequence-specific manner. These results indicate that CLM-MG exhibits many of the properties required for biologically useful, sequence-specific transcriptional antagonists. Unlike many other DNA ligands, the lipophilic nature of CLM-MG, in addition to favoring a minor groove interaction, allows for diffusion of the ligand through cell membranes, thereby permitting direct in vivo targeting of DNA–protein interactions. Furthermore, the sequence specificity and the inherent conformational rigidity of CLM-MG provide a means to specifically target and inhibit DNA–protein interactions. Greater sequence specificity in DNA–carbohydrate interactions should result from recent advances in carbohydrate chemistry permitting extended solid-phase polysaccharide synthesis (39) combined with DNA site-selection techniques. Thus, CLM-MG and other glycoconjugate DNA ligands may serve as lead compounds in the development of novel DNA-targeted therapeutic agents and biologic probes, perhaps ultimately leading to biologically useful DNA ligands.

This work was supported by grants from the National Institutes of Health to G.R.C. (CA 39612) and S.J.D. (HL 25848) and from the Howard Hughes Medical Institute to G.R.C. and by a Walter V. and Idun Y. Postdoctoral Fellowship to S.N.H.