Covalently Linking BHLH Subunits of MASH-1 Increases Specificity of DNA Binding[†]

Andreas G. E. Künne and Rudolf K. Allemann*

Department of Chemistry, ETH-Zürich, Universitätstrasse 16, CH-8092 Zurich, Switzerland

Received August 27, 1996; Revised Manuscript Received October 31, 1996[⊗]

ABSTRACT: MASH-1, a member of the basic-helix−loop−helix (BHLH) family of transcription factors, promotes the differentiation of committed neuronal precursor cells. *In vitro*, MASH-1 displays only marginal DNA sequence specificity. We have produced a MASH-1 variant, MASH−GGC, by introducing the tripeptide Gly-Gly-Cys at the C-terminal end of the BHLH domain. Under reducing conditions the properties of MASH−GGC and of the BHLH domain of MASH-1 were very similar. Like MASH-1, reduced MASH−GGC showed little specificity of DNA binding. CD spectroscopy revealed that both proteins underwent a conformational change from a largely unfolded to a mainly α-helical conformation upon binding to DNA. When the subunits of MASH−GGC were linked through a disulfide bond, the folded conformation was stable over a wide concentration range (2.5 nM to 2 μM) even in the absence of DNA. Oxidized MASH−GGC bound to E-box-containing sequences half-maximally at 148 nM, compared to 458 nM for the reduced form. Therefore, even when the change from a monomeric to a dimeric species was taken into account, the affinity for E-box-containing DNA sequences was increased. Surprisingly, the apparent dissociation constant for the complex with DNA not containing E-box sequences was increased upon oxidation. Therefore, despite the large distance between the disulfide bridge and the protein−DNA interface, covalently linking the subunits of MASH-1 increased the specificity of DNA binding significantly. *In vivo*, such an increase of the intrinsic DNA binding specificity might be achieved through interactions with other proteins of the transcriptional machinery.

Despite the high degree of sequence similarity between their DNA binding domains (Figure 1A) many of the members of the basic-helix−loop−helix (BHLH)[1] family of transcription factors are involved in the regulation of different biological processes. The expression of MyoD, for instance, can activate myogenesis in a wide variety of cell types including myoblasts and fibroblasts (Emerson, 1993; Lassar & Munsterberg, 1994; Olson, 1990), while MASH-1 promotes the differentiation of committed neuronal precursor cells (Sommer et al., 1995).

The activity of MASH-1 and of MyoD as transcriptional activators depends upon the presence of DNA sequences containing the symmetrical core motif CANNTG (E-box) (Figure 1C) (Lassar et al., 1989). MASH-1 and MyoD share a region of high sequence similarity called the BHLH domain, through which they dimerize and bind to DNA (Künne et al., 1996; Meierhans et al., 1995). Dimerization is mediated through the helix−loop−helix domain while the basic region of one BHLH subunit contacts one DNA half-site and the basic region of the other subunit interacts with the second half-site (Figure 1B) (Ma et al. 1994).

The binding preferences of MyoD and MASH-1 for different E-box sequences are similar, and the specificity of DNA binding is low (Künne et al., 1996; Meierhans et al., 1995). The apparent dissociation constants of MyoD and MASH-1 are only 1−2 orders of magnitude smaller for the complex with E-box-containing DNA sequences than for complexes with completely unrelated DNA sequences. Consequently, the specificity of transcriptional activation needed to explain the exquisite physiological specificity of MASH-1 and MyoD cannot be based solely on their intrinsic DNA binding specificities, but, most likely, will be achieved through cooperative interactions with other proteins of the transcriptional machinery such as the myocyte enhancer factors-2 (MEF-2) (Olson et al., 1995). While not myogenic themselves, MEF-2A and MEF-2C can act as coregulators to potentiate the myogenic effects of MyoD through interactions between the DNA binding domains (Molkentin et al., 1995). Similar interactions have been shown between MASH-1 and MEF-2A during the differentiation of the teratocarcinoma cell line P19 along a neuronal lineage (Mao & Nadal-Ginard, 1996).

When two transcription factors bind to DNA, the physical interaction between them can increase their individual DNA binding specificities by reducing the conformational flexibility of the amino acids at the protein−DNA interface, thereby either stabilizing the interaction with DNA sequences containing the preferred binding site or destabilizing the complex with heterologous DNA sequences or both.

As a model for such interactions, we constructed a MASH-1 BHLH variant, MASH−GGC, in which the tripeptide Gly-Gly-Cys was added to the carboxy-terminal end of the BHLH domain (Figure 1B). Under oxidizing conditions the subunits of MASH−GGC could be linked through a disulfide bond, thereby significantly increasing the stability of the folded form of the "dimer". Measurement of the

[†] This work was supported by an ETH grant (R.K.A.).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* January 1, 1997.

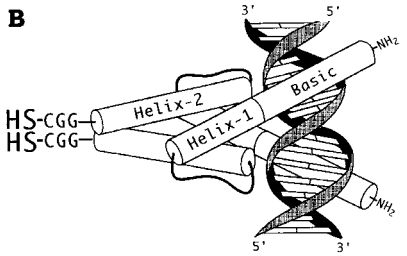
[1] Abbreviations: BHLH, basic-helix−loop−helix; bp, base pair(s); bZip, basic-region leucine zipper; CD, circular dichroism spectroscopy; EMSA, electrophoretic mobility shift assay; ^{1}H, proton; GGC-SH, reduced form of MASH−GGC; $(GGC\text{-}S)_2$, oxidized form of MASH−GGC; HPLC, high-performance liquid chromatography; IPTG, isopropyl β-D-1-thiogalctopyranoside; MEF-2, myocyte enhancer factors-2; MCK, muscle creatine kinase; PAGE, polyacrylamide gel electrophoresis; SD, standard deviation; TFE, 2,2,2-trifluoroethanol.

S0006-2960(96)02185-X CCC: $14.00

A

```
                          Basic       Helix-1      Loop        Helix-2
                        I----------I-----------I---------I----------------I
              110           120          130         140         150         160         170
MASH-GGC  GYSLPQQQPAA-----VARR---NERERNRVKLVNLGFATLREHVPNGAANKKMSKVETLRSAVEYIRALQQLLDGGC
MASH-1    GYSLPQQQPAA-----VARR---NERERNRVKLVNLGFATLREHVPNGAANKKMSKVETLRSAVEYIRALQQLLD
MASH-2     ATEASSSSAA-----VARR---NERERNRVKLVNLGFQALRQHVPHGGANKKLSKVETLRSAVEYIRALQRLLA
MyoD       CLLWACKACKRKTTNADRRKAATMRERRRLSKVNEAFETLKRCTSSNP-NQRLPKVEILRNAIRYIEGLQALLR
```

B



C

MCK-S	5'-CAGGCAG **CAGGTG** TTGG-3'
	3'-GTCCGTC **GTCCAC** AACC-5'
SP-1	5'-GATGCGGTCCCGCCCTCAGC-3'
	3'-CTACGCCAGGGCGGGAGTCG-5'

FIGURE 1: (A) Alignment of the BHLH domains of MASH−GGC, MASH-1, MASH-2, and MyoD. The full sequence of the MASH−GGC peptide used in this study is given. The tripeptide GGC, which was added to the C-terminal end of the BHLH domain of MASH-1, is underlined. The numbering corresponds to the full-length sequence of MASH-1. Amino acids are given in the one-letter code. Conserved residues are in bold face. The MASH-1 and MASH-2 sequences are from rat (Johnson et al., 1990). For MyoD the murine sequence is given (Davis et al., 1987). (B) Sketch of the complex of MASH−GGC with DNA. The tripeptide GGC is indicated, and the sulfhydryl groups are highlighted. The program Molscript (Kraulis, 1991) was used to make this display from the coordinates of the DNA complex of MyoD (Ma et al., 1994). α-Helices are represented as cylinders. The sugar−phosphate backbones of the DNA are represented by ribbons. (C) Sequences of the oligonucleotides used for the EMSAs. Both strands are shown. The E-box sequence is shown in bold face.

apparent dissociation constants of DNA complexes of MASH−GGC in its oxidized and its reduced form revealed that covalently linking the subunits reduced the K_D for the complex with an E-box-containing oligonucleotide almost 3-fold. On the other hand, the apparent dissociation constants for the complexes with heterologous DNA sequences were increased upon oxidation. Therefore, linking the subunits of MASH−GGC increased the DNA binding specificity by 1 order of magnitude despite the rather large distance between the disulfide bridge and the DNA−protein interface.

These results show that stabilizing the folded form of a BHLH protein by covalently linking the subunits through a disulfide bond can lead to increased DNA binding specificity. *In vivo*, such an increased specificity might result from interactions with other components of the transcriptional machinery, such as the myocyte enhancer factors MEF-2A and MEF-2C (Mao & Nadal-Ginard, 1996; Molkentin et al., 1995).

MATERIALS AND METHODS

Expression of MASH−GGC. MASH−GGC was expressed in BL21(DE3) cells containing the pLys plasmid (Studier, 1991) from the T7 promoter in the plasmid pJGetita (Meierhans et al., 1995). To construct the cDNA inserts of the expression plasmid, the rat MASH-1 BHLH sequence (Meierhans et al. 1995) was C-terminally elongated to encode the tripeptide, Gly-Gly-Cys, by inserting the nucleotide sequence GGAGGATGT followed by a TGA stop codon directly after the triplet encoding Asp(172) of MASH-1 through oligonucleotide-directed mutagenesis (Kunkel et al., 1987) with the oligonucleotide 5′-GCGCGCTGCAGCAGCT-GCTGGACGGAGGATGTTGATAGGATCCAC-TAGTTCTAGAGCGG-3′ using the Bio-Rad *in vitro* mutagenesis kit. Generally, the manufacturer's instructions were followed. The resulting cDNA construct contains the coding sequence for the amino acids Gly(106) to Asp(172) of MASH-1 and an additional tripeptide, Gly-Gly-Cys, at the C-terminus. The DNA sequence was verified using the dideoxy sequencing method (Sanger et al., 1977).

BL21(DE3) cells containing the MASH−GGC expression plasmid were grown at 37 °C on LB medium with 100 mg/L ampicillin and 50 mg/L chloramphenicol until the OD_{600} reached 0.4. Then IPTG was added to a final concentration of 1 mM. Cells were harvested 3 h after induction by centrifugation, and pellets were frozen at −20 °C.

Protein Purification. MASH−GGC was purified essentially as described for the BHLH region of MASH-1 (Meierhans et al., 1995). In short, the cells were resuspended in 3 mL of water/g of wet cells and 1 mM phenylmethanesulfonyl fluoride was added. After 2 vol of lysis buffer (100 mM ammonium acetate, pH 6.7, 100 mM sodium chloride, 100 mM 2-mercaptoethanol) had been added, cells were sonicated for 10 min on ice. The resulting suspension was dialyzed twice against urea buffer [5 mM sodium acetate (pH 5.0), 100 mM 2-mercaptoethanol, 8 M urea]. The dialyzate was centrifuged, and the supernatant was applied to a column containing 30 mL of Bio-Gel CM A ion-exchange resin (Bio-Rad) pre-equilibrated with urea buffer. The loaded resin was washed extensively with urea buffer and the protein eluted with one column volume of urea buffer containing 1 M sodium chloride. The eluate was dialyzed twice against urea buffer. The protein was further purified by preparative HPLC on a Resource-S sulfonate (Pharmacia) ion-exchange column (Meierhans et al., 1995). The collected fractions were pooled and concentrated by ultrafiltration using an Amicon YM-3 filter. The buffer was exchanged to 5 mM sodium acetate (pH 5.0) and 5 mM DTT by dialysis. SDS−PAGE showed a single protein band. MALDI-TOF showed a mass of 7817 which corresponded well with the calculated mass of 7819 for MASH−GGC without the N-terminal methionine (the accuracy of the measurement is 0.03%). The N-terminal sequence was confirmed by Edman degradation. The amino-terminal methionine had been removed proteolytically. Protein concentration was determined by measuring the UV absorption at 215 and 220 nm (Wetlaufer, 1962). The yields for the preparation were approximately 1 mg of purified protein/L of culture.

Dimerization of MASH−GGC. For dimerization a 40 μM solution of MASH−GGC was dialyzed against three changes of oxygen-saturated urea buffer not containing mercaptoethanol for 168 h followed by dialysis against two changes of 5 mM sodium acetate (pH 5.0). The protein concentration was determined by UV absorption (*vide supra*). SDS−PAGE showed a single protein band of about 15 kDa. MALDI-TOF showed sample homogeneity and a molecular mass of 15 640 ± 5 (calculated mass is 15 636).

Oligonucleotides. Oligonucleotides were purchased from Microsynth, desalted on Sephadex, and precipitated with ethanol. Single-stranded oligonucleotides were labeled with [^{32}P]ATP (Amersham) in the presence of T4 polynucleotide kinase (NEB), and complementary strands (10% excess) were annealed by heat denaturation followed by slow cooling to room temperature.

Electrophoretic Mobility Shift Assays. Aliquots from stock solutions of the proteins (monomer or dimer) were serially diluted (20 nM to 20 μM) into 5 mM sodium acetate (pH 5.0) and incubated in 50 mM Tris (pH 7.9), 6 mM $MgCl_2$, 40 mM $(NH_4)_2SO_4$, 0.2 mM EDTA, 4% (w/v) CHAPS, and 1 mM DTT for the reduced form of MASH−GGC. For oxidized MASH−GGC, DTT was omitted. After this mixture was incubated for 15 min at room temperature, 2 ng of labeled oligonucleotide probe was added and incubation continued for another 15 min at room temperature. The samples were applied to a 4% polyacrylamide gel in 1× TAE (pH 7.9). After electrophoresis the gels were dried and exposed to Kodak X-OMAT-S film at −70 °C. Quantitative data were obtained with an Instant Imager (Packard). The fraction Φ of DNA bound was determined as the activity of the retarded band (corresponding to the protein−DNA complex) divided by the sum of the activities of the retarded and unretarded (corresponding to the free DNA) bands.

CD Spectroscopy. Spectra were measured using Jasco J710 and J600 spectropolarimeters. The buffer was 5 mM Tris (pH 7.0) for oxidized MASH−GGC and 5 mM Tris (pH 7.0) containing 0.2 mM DTT for the reduced form of MASH−GGC. Spectra were measured for a concentration range from 2.5 nM to 2 μM for oxidized MASH−GGC and from 0.025 to 25 μM for the reduced form. For DNA binding experiments the protein concentration was 1 μM for the reduced form of MASH−GGC and 0.5 μM for the oxidized form.

RESULTS AND DISCUSSION

Design of Covalently Linked MASH−BHLH. A central feature of the reaction between BHLH domains and DNA is that protein folding, dimerization, and DNA binding are energetically interdependent processes (Meierhans et al., 1995; Sun & Baltimore, 1991). At high concentrations the BHLH domains form reasonably stable dimers with a K_{Dim} of approximately 4−30 μM (Figure 2C) (Fairman et al., 1993). However, at the concentrations where half-maximal DNA binding occurs (~10−100 nM), BHLH proteins exist as unfolded monomers in solution (Künne et al., 1996; Meierhans et al., 1995; Sun & Baltimore, 1991). Consequently, the free energy of DNA binding is reduced because at low protein concentrations energy must be spent on dimer formation. Therefore, covalently linking the subunits of a BHLH protein should obviate the requirement for dimerization and stabilize the folded form of the protein. Thus, such a protein should display enhanced DNA binding activity.

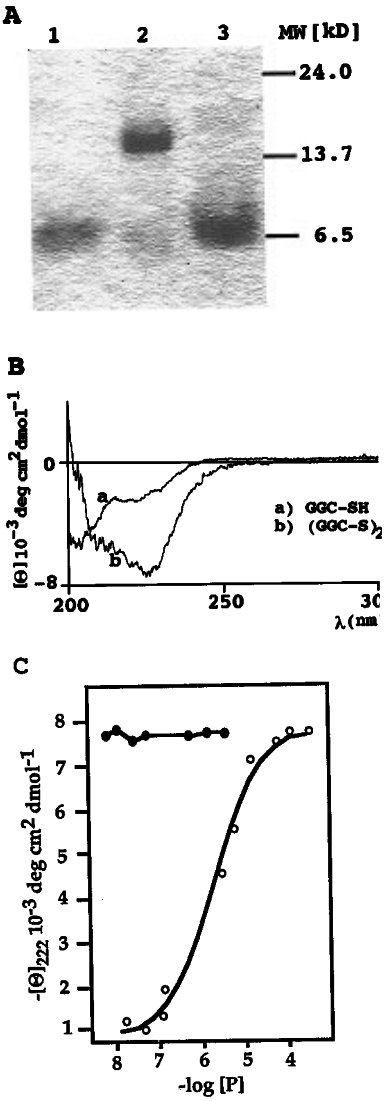


FIGURE 2: Characterization of MASH−GGC. (A) SDS−PAGE of purified proteins: lane 1, MASH-1 BHLH; lane 2, oxidized MASH−GGC; lane 3, reduced MASH−GGC. Mobilities of molecular weight markers (MW) are given in kDa. (B) CD spectra of reduced (a) and oxidized (b) MASH−GGC; [GGC-SH] = 1 μM; [(GGC-S)$_2$] = 0.5 μM. (C) Molar ellipticity [Θ] at 222 nm as a function of the concentration of oxidized MASH−GGC (filled circles) and MASH−BHLH (open circles). P refers to MASH−BHLH or (GGC-S)$_2$, respectively.

The X-ray structures of the DNA complexes of both MyoD (Ma et al., 1994) and E47 (Ellenberger et al., 1994) showed that the carboxy-terminal ends of their BHLH domains are in close proximity. We introduced the tripeptide, Gly-Gly-Cys, at the C-terminal end of the BHLH domain of MASH-1 to create the peptide MASH−GGC (Figure 1A and B). Simply changing the redox conditions can therefore induce a switch between the monomeric and the "dimeric" form, in which the subunits are held together via a disulfide bond. The glycines were introduced in order to increase the flexibility of the linker (Talanian et al., 1990) and to minimize the strain introduced by the disulfide bond.

Structural Characterization of MASH−GGC. MASH−GGC was produced in *Escherichia coli* and purified to apparent homogeneity. In denaturing polyacrylamide gel electrophoresis experiments the reduced form of MASH−GGC had similar mobility to MASH-1 BHLH (Figure 2A). Preparative oxidation of MASH−GGC in oxygen-saturated buffers was very slow, and the reaction did not reach completion after 7 days, possibly because of aggregation,

which is generally observed with BHLH proteins at high concentrations. However, in buffers containing 8 M urea the oxidation was essentially complete after 7 days and produced a single band of the correct molecular weight as judged by PAGE (Figure 2A) and mass spectroscopy.

CD spectroscopy was used to obtain further structural information about MASH−GGC. The BHLH domain of MASH-1 is largely unfolded at concentrations lower than 5 μM (Meierhans et al., 1995). Similarly, the CD spectrum of the reduced form of MASH−GGC showed that the peptide contained only small amounts of α-helicity at a concentration of 1 μM (Figure 2B). The oxidized form of MASH−GGC on the other hand showed significant amounts of α-helicity at a concentration of 0.5 μM (1 μM in monomer equivalents) (Figure 2B). The folded form of MASH-1 BHLH could be stabilized significantly by the addition of 2,2,2-trifluoroethanol (TFE) (Meierhans et al., 1995). Adding TFE to oxidized MASH−GGC did not result in any change in the CD spectrum, indicating that the protein was already fully folded. Unlike the non-linked forms of BHLH proteins (Figure 2C), which adopt stable secondary structures upon dimerization at protein concentrations between 4 and 30 μM (Fairman et al., 1993), oxidized MASH−GGC was clearly folded and mainly α-helical at these concentrations (Figure 2B). The mean residue ellipticity of oxidized MASH−GGC remained unchanged for the whole concentration range studied (2 μM to 2.5 nM) (Figure 2C). Therefore, the introduction of a disulfide linker between the subunits of MASH−GGC reduced significantly the concentration at which 50% of the protein molecules are in the folded form by at least 3 orders of magnitude relative to reduced MASH−GGC. This corresponds to a stabilization of the folded form of approximately 4.2 kcal/mol, an amount which is comparable to the stabilizations observed with single disulfide bonds engineered into globular proteins like T4 lysozyme (Matsumura et al., 1989). Consequently, oxidized MASH−GGC exists as a stably folded dimer at the protein concentration where half-maximal DNA binding occurs (Künne et al., 1996; Meierhans et al., 1995; Sun & Baltimore, 1991).

Structural Characterization of DNA Complexes of MASH−GGC. The DNA binding properties of MASH−GGC were analysed by electrophoretic mobility shift assay. As a DNA probe, an oligonucleotide (MCK-S) was chosen comprising 17 bp of the IgH enhancer-like element of the muscle creatine kinase enhancer (Buskin & Hauschka, 1989) with the central E-box sequence CAGGTG (Figure 1C). This sequence has been shown to be the highest affinity binding site for MASH-1 (Meierhans et al., 1995). For EMSAs, MASH−GGC was simply mixed with radioactively labeled MCK-S oligonucleotide (Figure 1C), and the products of the binding reaction were analyzed by native PAGE. Under these conditions, the DNA complexes of both the reduced and the oxidized forms of MASH−GGC produced retarded bands of equal electrophoretic mobility (Figure 3A), indicating that they had similar structures. In addition, the mobility shift was identical to that of the complex between MCK-S and the BHLH domain of MASH-1 (Figure 3A).

The DNA complexes of MASH−GGC were further characterized by CD spectroscopy. The addition of MCK-S oligonucleotide to a 1 μM solution of the reduced form of MASH−GGC induced a change in the CD spectrum indicative of a transition from a largely unfolded conformation of the peptide to a mainly α-helical form (Figure 3B) (Greenfield & Fasman, 1969). The same transition had been

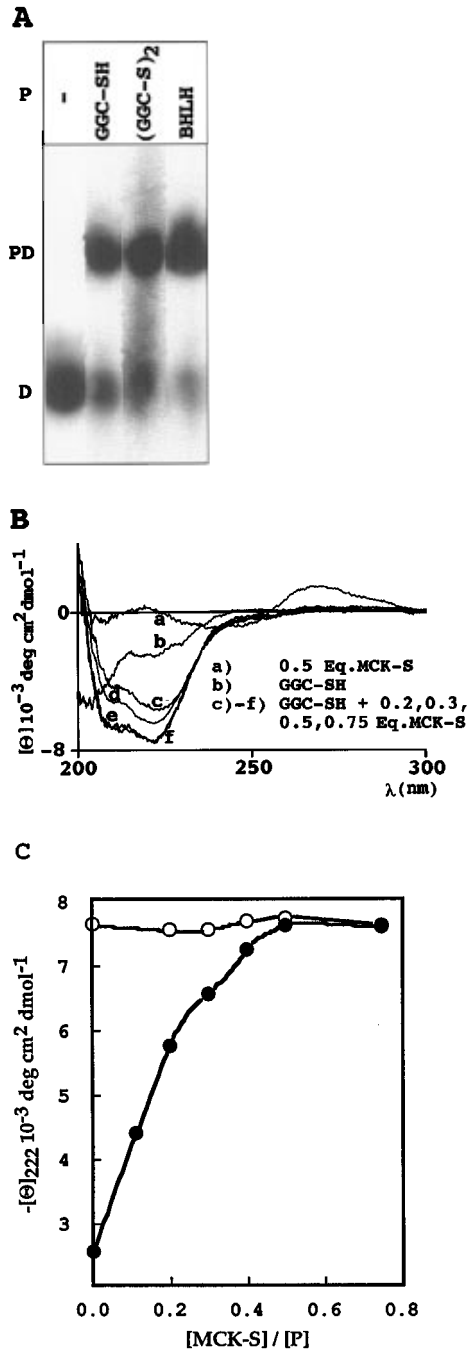


FIGURE 3: Characterization of the DNA complexes of MASH−GGC. (A) Autoradiogram of electrophoretic mobility shift assay with labeled MCK-S oligonucleotide and varying proteins. Lane 1 (−), MCK-S alone; lanes 2 and 3, labeled MCK-S with reduced (GGG-SH) and oxidized ((GGC-S)$_2$) MASH−GGC; lane 4, BHLH domain of MASH-1 (BHLH) (Meierhans et al, 1995). [MCK-S] = 9.66 nM; [GGC-SH] = 1.5 μM; [(GGC-S)2] = 1.0 μM; [BHLH] = 1.5 μM. (B) CD difference spectra of reduced MASH−GGC in the presence of increasing amounts of MCK-S oligonucleotide (the contribution from the oligonucleotide to the CD spectrum of the complex was subtracted). [MASH−GGC] = 1 μM. Curves b−f: [MCK-S]/[MASH−GGC] = 0, 0.2, 0.3, 0.5, 0.75; curve a, [MCK-S] = 0.5 μM (no protein present). (C) Molar ellipticity at 222 nm from CD difference spectra (the spectrum of the pure MCK-S oligonucleotide was subtracted from the spectrum of the complex) as a function of the ratio between the concentrations of MCK-S and reduced (filled circles) and oxidized (open circles) MASH−GGC, respectively. [GGC-SH] = 1 μM; [(GGC-S)$_2$] = 0.5 μM; [MCK-S] = 0, 0.2, 0.3, 0.4, 0.5, and 0.75 μM. Abbreviations: D, free MCK-S; P, protein; PD, protein−DNA complex.

observed with MASH-1 BHLH (Meierhans et al., 1995). The binding reaction could be monitored by measuring the

Table 1: DNA Binding Parameters for MASH-1 BHLH and Oxidized and Reduced MASH−GGC Determined by EMSA

	MCK-S[a]		SP-1[a]		
	K_D ($\times 10^{-13}$ M^2)[b]	$[P]_{1/2}$[c]	K_D ($\times 10^{-13}$ M^2)[b]	$[P]_{1/2}$[c]	specificity[d]
MASH-1	2.1 (±0.8)	458 nM	2.7 (±0.4)	520 nM	1.3
GGC-SH	2.1 (±0.5)	458 nM	2.5 (±0.2)	500 nM	1.2
$(GGC\text{-}S)_2$	0.9 (±0.1)	148 nM	9.9 (±0.4)	497 nM	11.0

[a] See Figure 1C for DNA sequences. [b] Apparent dissociation constants (±SD) for MASH−BHLH and the reduced form of MASH−GGC are reported as $K_D = ([P]_{1/2})^2$. For ease of comparison, K_D values (±SD) for oxidized MASH−GGC are reported per monomer equivalent [one molecule of $(GGC\text{-}S)_2$ contains two monomer equivalents] as $K_D = (2[(GGC\text{-}S)_2]_{1/2})^2$. [c] Concentration of protein for which 50% of the DNA binding sites are filled. [d] Specificity is calculated as the ratio K_D(SP-1)/K_D(MCK-S).

ellipticity at 222 nm (Figure 3C), which is indicative of the α-helical conformation. The change in the CD spectrum upon addition of DNA could be saturated, in that the addition of excess oligonucleotide did not result in a further change of the spectrum (Figure 3B and C). The CD spectrum of the complex between MCK-S oligonucleotide and reduced MASH−GGC (Figure 3B) was nearly identical to the spectrum of oxidized MASH−GGC in the absence of DNA (Figure 2B), providing further evidence that the folded forms of DNA free, oxidized MASH−GGC and of reduced MASH−GGC bound to DNA were very similar. The addition of oligonucleotide to the oxidized form of MASH−GGC did not change the CD spectrum, and the ellipticity at 222 nm did, therefore, not increase further (Figure 3C).

The results described above suggested that linking the subunits of MASH−GGC through a disulfide bond uncoupled the processes of folding, dimerization, and DNA binding, because at the concentration where DNA binding occurs the oxidized form of MASH−GGC is already folded into a stable covalent "dimer".

DNA Binding Specificity of MASH−GGC. In EMSA titrations, the apparent dissociation constants (K_D) were measured for the complexes of the reduced and the oxidized forms of MASH−GGC with the MCK-S oligonucleotide (Figure 1C) and with a completely unrelated DNA sequence containing a binding site for the transcription factor SP-1 (Kadonaga & Tijan, 1986) (Figure 1C). It is known that the BHLH region of MASH-1 shows little DNA binding specificity and binds to MCK-S only 1.3 times tighter than to unrelated DNA sequences (Table 1) (Meierhans et al., 1995). The concentration of the oligonucleotide in these experiments was held constant while the concentration of the proteins was successively increased (Figures 4A and B). For the reduced form of MASH−GGC, the best fit for the binding isotherm (eq 1) was found for the assumption of cooperative binding of two monomers of the reduced form of MASH−GGC to DNA (Figure 4C), while for oxidized MASH−GGC the best fit for the isotherm (eq 2) was observed under the assumption of binding of one monomer (corresponding to two BHLH chains) to the DNA (Figure 4D):

$$\Phi = K^2[P]^2/(1 + K^2[P]^2) \quad (1)$$

$$\Phi = K[P]/(1 + K[P]) \quad (2)$$

From the isotherms, the protein concentration for half-maximal binding $[P]_{1/2} = 1/K$ could be determined. To facilitate comparison of the data, apparent dissociation constants were calculated relative to monomer equivalents (one molecule of oxidized MASH−GGC contains two monomer equivalents, while one molecule of the reduced form contains one monomer equivalent) as $K_D = ([P]_{1/2})^2$ for the reduced form and $K_D = (2[P]_{1/2})^2$ for the oxidized form of MASH−GGC.

The reduced form of MASH−GGC showed approximately the same affinity for the oligonucleotides MCK-S and SP-1 (Table 1; Figure 4A−C). The apparent dissociation constants were similar to the K_D values measured for the complexes of the unmodified BHLH region of MASH-1 with these oligonucleotides (Table 1) (Meierhans et al., 1995).

The results obtained for the oxidized form of MASH−GGC were dramatic in that the amount of stabilization of the protein−DNA complex through the introduction of the disulfide bond between the BHLH subunits was dependent on the specific DNA sequence. Even when the change from a monomeric to dimeric species was taken into account, oxidized MASH−GGC bound E-box-containing DNA sequences at smaller half-maximal concentrations than the reduced form (Table 1, Figure 4A−D). The apparent dissociation constants were 0.9×10^{-13} M^2 for oxidized and 2.1×10^{-13} M^2 for reduced MASH−GGC. On the other hand, approximately equal concentrations of oxidized and reduced MASH−GGC were needed for half-maximal binding to non-specific DNA (Table 1, Figure 4B−D). The binding affinity of each BHLH chain to heterologous DNA was reduced 2-fold in the oxidized form of MASH−GGC compared to the non-covalent dimer formed by reduced MASH−GGC. Consequently, the apparent dissociation constant for the complex of non-specific DNA was almost four times bigger with oxidized MASH−GGC than with the reduced form (Table 1). Overall the stabilization of the complex with E-box-containing DNA and the destabilization of the complex with non-specific DNA upon oxidation resulted in an increase of the DNA binding specificity of MASH−GGC of 1 order of magnitude (Table 1).

Because of the large distance between the disulfide bridge and the basic region of MASH−GGC, and because the results from CD spectroscopy suggested that the DNA complexes of the oxidized and the reduced forms of MASH−GGC had similar structures, it was unlikely that the 10-fold increase in DNA binding specificity was simply the consequence of either additional favorable contacts between residues of the linker region and the nucleotide bases of the E-box containing DNA or new unfavorable interactions with the SP-1 oligonucleotide.

Covalently linking the subunits of MASH−GGC through a disulfide bond might reduce the number of conformations accessible to the protein in the disordered state and, therefore, diminish the entropic penalty that accompanies folding and DNA binding. Although the flexibility of side chains at the protein−DNA interface is restrained when compared to the uncomplexed state, it is still higher than the flexibility of side chains in the protein core (Berglund et al., 1995). Stabilizing the folded form of MASH−GGC through a disulfide bond might compensate for an entropically unfavorable reduction of the conformational flexibility of the side chains at the protein−DNA interface. Limiting the number of accessible conformations of the basic region of BHLH proteins could stabilize the complex with specific DNA and destabilize the complex with non-specific DNA sequences, thereby increasing the overall DNA binding specificity. A precedent for such a mechanism is provided by the tryp-

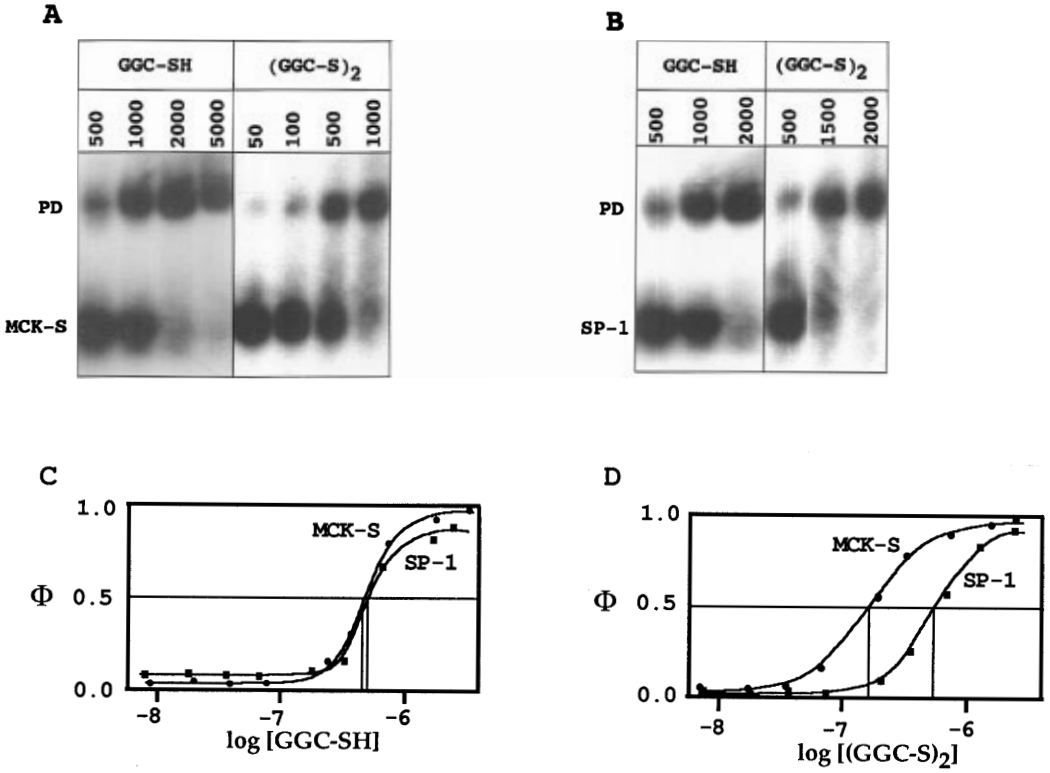


FIGURE 4: (A) Autoradiograms of electrophoretic mobility shift assay titrations of radiolabeled MCK-S oligonucleotide with increasing amounts of reduced (left panel) and oxidized (right panel) MASH−GGC. [MCK-S] = 9.66 nM; protein concentrations are given in nM. (B) As in A, but SP-1 oligonucleotide was used; [SP-1] = 8.2 nM. (C) Fraction Φ of bound DNA in the complex of reduced MASH−GGC with oligonucleotides MCK-S and SP-1, respectively, as a function of the concentration of reduced MASH−GGC. The sigmoidal lines represent the best theoretical fit through the data points. The vertical lines indicate the concentrations of protein, where half of the protein binding sites on the DNA are filled. (D) As in C, but the oxidized form of MASH−GGC was used.

tophan repressor of *E. coli*. Replacing alanine (77) with valine, for example, results in a local stabilization of the DNA binding domain of the repressor, reduced conformational flexibility (Gryk & Jardetzky, 1996), and increased specificity through reduced affinity for non-specific DNA (Arvidson et al., 1993). In this context, it is worthy of note that the addition of 20% TFE to the BHLH domain of MASH-1 increased not only the DNA binding specificity by approximately 2-fold (Meierhans et al., 1995) but also significantly increased the dispersion of the ^{1}H-NMR signals, suggesting a decrease of the conformational flexibility of the protein (data not shown).

Because the milieu inside a cell is strongly reducing, the conformational rigidity of the basic region of MASH-1 and the specificity of DNA binding cannot be increased *in vivo* by covalently linking the subunits through a disulfide bond. However, such an increase of the specificity might result from the interaction with other components of the transcriptional machinery, for instance with MEF-2A (Mao & Nadal-Ginard, 1996). In this case, MEF-2A could stabilize a specific conformation of the basic region of MASH-1 and thereby alter the intrinsic DNA binding specificity of MASH-1. A direct interaction of MEF-2A with the α-helix of the basic domain of MASH-1 might lead to a significantly increased effect when compared to the stabilization through a remote disulfide bond (Figure 1B). Such a mechanism could resolve the discrepancy between the low intrinsic DNA binding specificity and the exquisite physiological specificity of BHLH proteins.

A similar mechanism has been described previously for the human T-cell leukaemia virus type-1 Tax protein, which binds to the basic region of basic-zipper proteins (bZip) (Baranger et al., 1995; Perini et al., 1995). When present in large excess, Tax enhances the affinity of bZip proteins for cyclic AMP-responsive element target sites (Armstrong et al., 1993; Wagner & Green, 1993). Surprisingly and consistent with the model presented above, binding of Tax to various bZip proteins also changed their relative affinities for different DNA binding sites (Perini et al., 1995).

The work described in this paper represents the first instance in which a significant increase in the DNA binding specificity of a transcription factor has been achieved by covalently cross-linking its subunits through a disulfide bond. It provides opportunities to address questions about the molecular mechanisms by which the DNA binding specificity of transcription factors can be modulated.

ACKNOWLEDGMENT

We thank Steven A. Benner for his support and the use of his laboratory; Joseph Stackhouse for providing part of Figure 2C; David J. Anderson for providing the MASH-1 cDNA clone; Hans Ulrich Hediger for MALDI-TOF mass spectroscopy; Steven A. Benner and Lesley A. Tannahill for critical reading of the manuscript; and the members of our laboratory for helpful discussions.

REFERENCES

Armstrong, A. P., Franklin, A. A., Uittenbogaard, M. N., Giebler, H. A., & Nyborg, J. K. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 7303−7307.

Arvidson, D. N.,. Pfau, J., Hatt, J. K., Shapiro, M., Pecoraro, F. S., & Youderian, P. (1993) *J. Biol. Chem. 268*, 4362−4369.

Baranger, A. M., Palmer, C. R., Hamm, M. K., Giebler, H. A., Brauweiler, A., Nyborg, J. K., & Schepartz, A. (1995) *Nature 376*, 606−608.

Berglund, H., Baumann, H., Knapp, S., Ladenstein, R., & Härd, T. (1995) *J. Am. Chem. Soc. 117*, 12883−12884.
Buskin, J. N., & Hauschka, S. D. (1989) *Mol. Cell. Biol. 9*, 2627−2640.
Davis, R. L., Weintraub, H., & Lassar, A. B. (1987) *Cell 51*, 987−1000.
Ellenberger, T., Fass, D., Arnaud, M., & Harrison, S. C. (1994) *Genes Dev. 8*, 970−980.
Emerson, C. P. (1993) *Curr. Opin. Genet. Dev. 3*, 265−274.
Fairman, R., Beran-Steed, R. K., Anthony-Cahill, S. J., Lear, J. D., Stafford, W. F., III, DeGrado, W. F., Benfield, P. A., & Brenner, S. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 10429−10433.
Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108−4115.
Gryk, M. R., & Jardetzky, O. (1996) *J. Mol. Biol. 255*, 204−214.
Johnson, J. E., Birren, S. J., & Anderson, D. J. (1990) *Nature 346*, 858−861.
Kadonaga, J. T., & Tijan, R. (1986) *Cell 51*, 1079−1090.
Kraulis, R. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
Künne, A. G. E., Meierhans, D., & Allemann, R. K. (1996) *FEBS Lett. 391*, 79−83.
Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., & Weintraub, H. (1989) *Cell 58*, 823−831.
Lassar, A. B., & Munsterberg, A. (1994) *Curr. Opin. Cell. Biol. 6*, 432−442.
Ma, P. C. M., Rould, M. A., Weintraub, H., and Pabo, C. O. (1994) *Cell 77*, 451−459.
Mao, Z., & Nadal-Ginard, B. (1996) *J. Biol. Chem. 271*, 14371−14375.
Matsumura, M., Signor, G., & Matthews, B. W. (1989) *Nature 342*, 291−293.
Meierhans, D., el Ariss, C., Neuenschwander, M., Sieber, M., Stackhouse, J. F., & Allemann, R. K. (1995) *Biochemistry 34*, 11026−11036.
Molkentin, J. D., Black, B. L., Martin, J. F., & Olson, E. N (1995) *Cell 83*, 1125−1136.
Olson, E. N. (1990) *Genes Dev. 4*, 1454−1461.
Olson, E. N., Perry, M., & Schulz, R. A. (1995) *Dev. Biol. 172*, 2−14.
Perini, G., Wagner, S., & Green, M. R. (1995) *Nature 376*, 602−605.
Sanger, F., Nicklen, S., Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5469.
Sommer, L., Shah, N., Rao, M., & Anderson, D. J. (1995) *Neuron 15*, 1245−1258.
Studier, F. W. (1991) *J. Mol. Biol. 219*, 37−44.
Sun, X.-H., & Baltimore, D. (1991) *Cell 64*, 459−470.
Talanian, R. V., McKnight, C. J., & Kim, P. S. (1990) *Science 249*, 769−771.
Wagner, S., & Green, M. R. (1993) *Science 262*, 395−399.
Wetlaufer, D. R. (1962) *Adv. Prot. Chem. 17*, 303−390.

BI962185L