Doctoral Thesis

DNA binding reaction of basic-helix-loop-helix proteins

Author(s):
Sieber, Martin

Publication Date:
2000

Permanent Link:
https://doi.org/10.3929/ethz-a-003913547

Rights / License:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
DNA Binding Reaction of
Basic-Helix-Loop-Helix Proteins

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Natural Sciences

presented by
Martin Sieber
Dipl. Chem. ETH
born January 3, 1970
citizen of Beinwil (SO)

accepted on the recommendation of

Prof. Dr. Donald Hilvert, examiner
Dr. Rudolf K. Allemann, co-examiner

Zurich 2000
Acknowledgement

I would like to thank Dr. Rudolf K. Allemann for the possibility to experience a scientific environment at the ETH Zurich and at the University of Birmingham and Prof. Dr. D. Hilvert for being my examiner.

A very special thank belongs to Yi Yang who believed in me and supported me even from far away.
And a thanks to all my friends for their support. Especially Daniel Meierhans, Andreas Künne, Andrea Piatesi, Egle Bieliauskaite, Barbara and Edith Schielly, Susan Taylor, Libby Turner and Jon Moore.
Table of contents

Table of contents ............................................................... 1

Abbreviations ................................................................. 5

Abbreviations for amino acids ............................................. 7

Summary .............................................................................. 8

Zusammenfassung .............................................................. 11

Chapter 1: Introduction ....................................................... 14

Chapter 2: Alteration of the binding specificity of E12 ........ 20

2.1 Introduction ..................................................................... 20

2.2 Rational design of a E12-DNA complex with new specificity ... 21

2.3 Construction of the E12 mutants ..................................... 25

2.4 Expression and Purification of E12 and mutants ................. 28
2.5 Mass spectrometry of E12 and mutants ................................................. 31

2.6 Binding affinities of the E12 mutants ..................................................... 33

Chapter 3: A Single Chain 'Dimer' of MASH-BHLH .................. 40

3.1 Design of a bHLH single chain 'dimer' ................................................. 40

3.2 Construction of MM8 and MM17 ............................................................ 42

3.3 Construction of ME8 and ME17 ............................................................. 46

3.4 Expression of MM17 ............................................................................. 51

3.5 Purification of MM17 ............................................................................ 52

3.6 CD spectroscopy of MM8 and MM17 .................................................... 56

3.7 DNA binding affinity of MM8 and MM17 .............................................. 62

3.8 Specificity of DNA binding ................................................................. 66
Chapter 4: Thermodynamics of MASH BHLH/DNA complex formation

4.1 Introduction ......................................................... 68

4.2 pH dependence of the binding reaction ................................ 68

4.3 ITC measurements of the binding reaction ........................... 70

4.4 Calculation of the thermodynamic parameters ...................... 76

4.5 Thermodynamic parameters for the binding process .............. 79

4.6 Coupling of MASH-BHLH folding to DNA binding ................ 82

4.7 CD spectroscopy studies of the binding process ................... 84

Chapter 5: Thermodynamics of MM17/DNA complex formation

5.1 Introduction .................................................................. 88

5.2 ITC Measurements ..................................................... 89

5.3 pH dependence of the binding process ............................... 93
5.4 Thermodynamic parameters for the binding process ..................... 94

5.5 Interpretation of the observed heat capacity change ..................... 95

Chapter 6. Conclusions .................................................................. 99

Chapter 7: Materials and Methods ................................................. 101

7.1 Materials .................................................................................. 101

7.2 Methods .................................................................................... 104

References ...................................................................................... 128

Curriculum Vitae ............................................................................ 137
Abbreviations

Amp Ampicillin
AP Ammoniumperoxodisulfate
BHLH Basic Helix Loop Helix
BHLHZ Basic Helix Loop Helix Zipper
CD Circular Dichroism
Cfu Colony forming units
Cl-amp Chloramphenicol
DTT Dithiothreitol
EMSA Electrophoretic Mobility Shift Assay
HPLC High Performance Liquid Chromatography
Ig H Immunoglobulin Heavy Chain
IPTG Isopropyl-β-D-thiogalactopyranoside
ITC Isothermal Titration Calorimetry
LB Luria Bertani
MALDI-TOF Matrix Assisted Laser Desorption Ionisation - Time of Flight
MASH Mammalian Achaete Scute Homologue
MCK Muscle Creatine Enhancer
MES 2-Morpholinoethanesulfonic acid
MOPS 3-Morpholinopropanesulfonic acid
MW Molecular Weight
OD Optical Density
PAGE Polyacrylamid Gel Electrophoresis
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pfu</td>
<td>Plaque forming units</td>
</tr>
<tr>
<td>PIPES</td>
<td>Piperazine-1,4-bis(2-ethanesulfonic acid)</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethanesulfonyl fluoride</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecylsulfate</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N',N'-Tetramethylethylenediamine</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)-aminomethane</td>
</tr>
<tr>
<td>Max</td>
<td>Myc-associated x protein</td>
</tr>
<tr>
<td>USF</td>
<td>Upstream Stimulatory Factor</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>Id</td>
<td>Inhibitor of differentiation protein</td>
</tr>
<tr>
<td>Hes</td>
<td>Hairy- and Enhancer of split-related mammalian proteins</td>
</tr>
<tr>
<td>Myf</td>
<td>Myogenic Factor</td>
</tr>
<tr>
<td>MRF4</td>
<td>Muscle Regulatory Factor 4</td>
</tr>
<tr>
<td>MyoD</td>
<td>Myogenic Determination Factor</td>
</tr>
</tbody>
</table>
## Abbreviations for amino acids

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Abbreviation</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>Q</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
</tr>
<tr>
<td>Tryptophane</td>
<td>Trp</td>
<td>W</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
</tr>
</tbody>
</table>
Summary

The initiation of transcription of genes within eukaryotic cells is controlled by complex interactions between transcription factors and specific DNA recognition sequences in target genes. Therefore, transcription factors are critical regulators of gene expression. Basic helix-loop-helix proteins are one group of transcription factors. They consist of an HLH motif, which mediates dimerisation between HLH proteins, and a basic region which binds DNA through interactions between positively charged residues and the negatively charged DNA backbone and by making specific contacts to the nucleobases. BHLH proteins usually recognise the palindromic DNA sequence CANNTG, called the E-box.

In order to understand the molecular basis for site specific DNA recognition of BHLH proteins, several mutants of the protein E12 were produced. The dissociation constants were measured by EMSAs with a range of oligonucleotides. In wild-type E12, two residues are important for site-specific target recognition: Glu 345, which accepts hydrogen bonds from N4 of cytosine and N6 of adenine, and Arg 348, which fixes the position of Glu 345 by contacting the phosphate backbone and forming a salt bridge to Glu 345. The $K_D$ of the complex of E12 with the E-box containing oligonucleotide MCK-S was $2.7 \pm 0.55 \times 10^{-15}$ M$^2$ and the specificity to heterologous DNA approximately 100. Binding to MCK-SQ (CGGGCG) was weaker ($26.2 \pm 10.0 \times 10^{-15}$ M$^2$) presumably because of a repulsive interaction between O4 of guanine and the carboxylate oxygens of Glu 345. Replacing the glutamate with a glutamine produced the mutant E345Q which formed a more stable complex with MCK-SQ than E12, because a similar set of contacts like in the
wild-type complex could be formed. It also bound heterologous oligonucleotides with enhanced affinity. The reason for the high affinity could be the missing charge at position 345 so that Arg 348 can move closer to the DNA backbone. The mutant E345L showed little specificity but the binding affinity was reduced by one order of magnitude. Changing Arg 348 to lysine reduced the affinity dramatically. Because lysine cannot form a 'clamp', it presumably moves either toward the phosphate or Glu 345.

The two subunits of MASH-BHLH or E12-BHLH were connected through linkers of different length from the C-terminus of one BHLH protein to the N-terminus of the other producing the single chain dimers MM8, MM17 and ME17. CD spectroscopy revealed that approximately 45% of the protein was in α-helical conformation. Upon addition of 1 equivalent of DNA a folding transition to a mainly α-helical conformation was observed. The single chain dimers behaved similarly to wild-type MASH-BHLH. However, a 10 fold increase in the binding affinities compared to wild-type MASH-BHLH was observed.

Thermodynamic parameters for the binding process of MASH-BHLH to the MCK-S and SP-1 oligonucleotides were determined by isothermal titration experiments. The association constants for the binding reaction between MASH-BHLH and MCK-S and SP-1 were 3.41 (±0.1) x 10^7 M⁻¹ and 2.08 (±0.3) x 10^7 M⁻¹ at 12°C and 51.25 (±5.1) x 10^7 M⁻¹ and 5.14 (±1.9) x 10^7 M⁻¹ at 27°C. The enthalpy and the entropy of binding were strongly dependent on temperature. As a result, the free energy of binding was almost temperature independent. The association process was enthalpically driven throughout the physiological temperature range and characterised by a large negative heat capacity change of -733 (±99) cal mol⁻¹ K⁻¹ for MCK-S and -575 (±105) cal
mol$^{-1}$ K$^{-1}$ for SP-1. Within experimental error, the association was independent of pH between pH 6 and 8. Dissection of the entropy change into its component parts indicated that binding was coupled to local protein folding of approximately 25 amino acids per protein subunit, an observation which was consistent with CD measurements. Therefore, the basic region undergoes a folding transition upon DNA binding.

To complete the biochemical and physical analysis of the properties of MM17, ITC measurements were performed to obtain thermodynamic parameters of the DNA binding process. As for MASH-BHLH, the DNA binding of MM17 was not pH-dependent in the range between pH 6 and 8. The association was enthalpically driven throughout the physiological temperature range examined. The association constants for MM17 binding to MCK-S and SP-1 were $1.11 \pm 0.1 \times 10^7$ M$^{-1}$ and $5.04 \pm 2.75 \times 10^6$ M$^{-1}$ at $12^\circ$C and $4.38 \pm 1.8 \times 10^7$ M$^{-1}$ and $6.58 \pm 2.9 \times 10^6$ M$^{-1}$ at $32^\circ$C, respectively. Interpretation of the observed heat capacity changes and dissection of the entropy changes revealed that approximately 70 amino acids adopt a folded form on DNA binding. In addition to the 54 residues of the basic regions, 16 amino acids lose conformational degrees of freedom in the complex which could mean the linker in MM17 adopts a well defined structure upon complex formation.
Zusammenfassung


Um die molekulare Grundlage für spezifische DNA Erkennung zu verstehen, sind verschiedene Mutanten des Proteins E12 hergestellt worden. Die Dissoziationskonstanten mit verschiedenen Oligonukleotiden wurden mit EMSAs gemessen. In wild-typ E12 sind zwei Aminosäuren für die sequenzspezifische Erkennung wichtig: Glu 345 macht eine Wasserstoffbrücke zu N4 von Cytosin und N6 von Adenin; Arg 348 fixiert die Position von Glu 345, indem zugleich das Phosphatrückgrad kontaktiert und eine Salzbrücke zu Glu 345 ausgebildet wird.

Der K\textsubscript{D} des E12-Komplexes mit dem Oligonukleotid MCK-S war 2.7 (±0.55) x 10\textsuperscript{-15} M\textsuperscript{2} und die Spezifität ungefähr 100. Die Affinität zu MCK-SQ (CGGGCG) wurde geschwächt (26.2 (±10.0) x 10\textsuperscript{-15} M\textsuperscript{2}) wegen einer abstossenden Wechselwirkung zwischen O4 von Guanin und den Carboxylat Sauerstoffen von Glu 345. Wird Glutamat ersetzt durch Glutamin, erhält man


Die thermodynamischen Parameter für die Binding von MASH-BHLH mit den Oligonukleotiden MCK-S und SP-1 wurden durch isothermische Titrations-experimente bestimmt. Die Assoziationskonstanten zwischen MASH-BHLH und MCK-S bzw. SP-1 waren 3.41 (±0.1) x 10^7 M^-1 und 2.08 (±0.3) x 10^7 M^-1 für 12°C und 51.25 (±5.1) x 10^7 M^-1 und 5.14 (±1.9) x 10^7 M^-1 für 27°C. Die Assoziations-enthalpie und -entropie waren stark Temperatur abhängig. Dadurch wurde die freie Bindungs-energie praktisch Temperatur abhängig. 
unabhängig. Die Assoziation war im ganzen physiologischen Temperaturbereich enthalpisch getrieben und wurde durch eine stark negative Änderung der Wärmekapazität von -733 (±99) cal mol⁻¹ K⁻¹ für MCK-S und -575 (±105) cal mol⁻¹ K⁻¹ für SP-1 gekennzeichnet. Innerhalb des experimentellen Fehlers war die Assoziation im Bereich von pH 6 bis 8 unabhängig vom pH. Zerlegung der Änderung der Entropie zeigte, dass die Bindung zu lokaler Proteinfaltung von ungefähr 25 Aminosäuren pro Einheit gekoppelt war. Diese Beobachtung war konsistent mit CD Messungen. Das zeigt, dass die basische Region einen Faltungsübergang bei der Bindung von DNA durchläuft.

Um die biochemische und physikalische Analyse der Eigenschaften von MM17 zu vervollständigen, wurden ITC Messungen durchgeführt und dadurch die thermodynamischen Parameter der Bindungsreaktion erhalten. Wie für MASH-BHLH war die DNA Bindung von MM17 im Bereich von pH 6 bis 8 nicht pH abhängig. Ebenfalls war die Assoziation über den gemessenen physiologischen Temperaturbereich enthalpisch getrieben. Die Assoziationskonstanten der Bindung von MM17 mit MCK-S bzw. SP-1 waren 1.11 (±0.1) x 10⁷ M⁻¹ und 5.04 (±2.75) x 10⁶ M⁻¹ für 12°C und 4.38 (±1.8) x 10⁷ M⁻¹ und 6.58 (±2.9) x 10⁶ M⁻¹ für 32°C. Interpretation der Wärmekapazitätsänderung und Zerlegung der Änderung der Assoziationsentropie zeigte, dass etwa 70 Aminosäuren auf Grund von der DNA Bindung eine gefaltete Konformation annehmen. Zusätzlich zu den 54 Resten der basischen Regionen, verlieren 16 Aminosäuren konformationelle Freiheitsgrade im Komplex, was heissen könnte, dass der Linker in MM17 eine definierte Struktur annimmt.
Chapter 1: Introduction

In billions of years nature has evolved a system with countless different lifeforms. This system is based on the interactions of the parts within it. Complex relationships between a vast array of animals, plants, fungi, insects and micro-organisms are what keeps the biosphere alive. The importance of every single lifeform is realised as soon as we consider that if only one type of organism is missing and hence doesn't fulfil its function, a catastrophic chain reaction could be the consequence followed by the breakdown of the system. We find the necessity of interactions for functionality of the system on every level down to individual molecules. This amazing interplay results in the wonderful diversity of living organisms. What is hidden before our eyes on the cellular level is even more interesting and it is a challenge to find an explanation for it. Albert Einstein once said: "The important thing is not to stop questioning. Curiosity has its own reason for existing". Therefore exploring the world of biological chemistry will reveal the answers to the questions about life.

In order to explain how nature has produced all this, it is vital to understand the mechanisms that make life possible. Although molecules are not visible to the naked eye, the effects of their interactions are observable and constitute the basis of natural sciences.

DNA is the carrier of the genetic code that contains all the information needed to grow an organism and to keep it alive. A major goal is to understand the interactions involved in regulating the development of an individual. The fundamental dogma of molecular biology is that DNA produces RNA, which
in turn produces protein. The process of transcription, whereby an RNA product is produced from a DNA template, is an essential element in gene expression. The failure of this process to occur will obviously render redundant all the other steps that follow the production of the initial RNA transcripts, such as translation into protein. The central role of transcription in the process of gene expression is also an attractive control point for regulating the expression of genes.

In eukaryotes the production of RNA from a DNA template is catalysed by DNA-dependent RNA polymerases. RNA polymerase and additional proteins (basal transcription factors) build up the basal transcription machinery, which is necessary for low level transcriptional initiation from the proper start point. As biological processes are complex, it makes it obvious that gene expression is tightly controlled. For the initiation of RNA synthesis, this is achieved through the action of transcription factors. These proteins bind with high affinity to their target DNA sequences and regulate the rate of transcription through interaction with the basal transcriptional machinery. The specificity of transcriptional initiation is controlled through the interaction of transcription factors with their DNA-binding sites in the promoter and enhancer regions of genes. Hence it is vital to understand the forces that govern the recognition of DNA by proteins.

Nature has used a modular approach to assemble transcription factors. The DNA-binding domain and the transcriptional activation domain are therefore to a first approximation independent of each other. As a consequence, the properties of the DNA-binding domains can be studied in the absence of the activation domain.
Most DNA-binding proteins can be grouped into families based on the structural motif that they rely on for sequence specific DNA recognition. Several DNA binding motifs have been characterised such as the helix-turn helix motif, the MADS Box, the leucine zipper, the homeo domain and zinc fingers. Another motif is the helix-loop-helix (HLH) domain, first identified in an immunoglobulin enhancer-binding polypeptide (Murre et al., 1989). It is present in many proteins of diverse biological function unified only by their common involvement in transcriptional regulation (Amati & Land, 1994; Jan & Jan, 1993; Kadesch, 1992; Li & Olson, 1992; Olson, 1993; Tapscott et al., 1990; Wright, 1992). The HLH domain is usually adjacent to a short region of basic residues and therefore these proteins are referred to as bHLH proteins. A second group of bHLH proteins contain an additional dimerisation motif, the leucine zipper (Landschulz et al., 1988), immediately C-terminal to the HLH. Examples are Myc, Max or USF (Bernard et al., 1983; Blackwood & Eisenman, 1991; Gregor et. Al., 1990). These proteins are commonly referred to as bHLHZ proteins. Members of a third group of HLH proteins such as Id or Hes lack a functional DNA-binding domain and act as negative regulators of bHLH proteins (Sun et al., 1991; Riechmann et al., 1994; Benezra et al., 1990; Sasai et al., 1992; Ishibashi et al., 1993).

BHLH proteins can be divided into two broad functional groups by their patterns of expression. In the first group, expression is restricted to cells of a particular lineage. Examples are the myogenic factors MyoD, Myf5, MRF4, myogenin (Braun et al., 1989; Pearson, 1991) or the neurogenic factors MASH-1 and MASH-2 (Johnson et al., 1990). The second group are expressed fairly ubiquitously. These proteins, like E12 and E47, typically form heterodimers with the lineage-specific bHLH proteins.
BHLH proteins consist of two helices separated by a non-conserved sequence of variable length (the loop) that mediate homo-/hetero- dimerisation with other bHLH proteins. N-terminal to helix 1 follows a short stretch of hydrophilic (often basic) residues that mediates DNA binding (Figure 1.1). These proteins bind DNA as dimers and usually recognise a short palindromic sequence, named E-box (CANNTG), such that each protein monomer binds on one half-site. Sequence specificity is assumed to reside in the combination of the two basic domains in the dimer.

BHLH dimers form a parallel, left-handed, four-helix bundle with a stable hydrophobic core. Helix 1 packs against helix 2 in the same molecule and also against helix 2' of the opposing subunit. Extensive van der Waals contacts and limited electrostatic interactions stabilise the HLH four-helix bundle in the dimer.

**Figure 1.1:** Sketch of a BHLH-DNA complex (From Ma et al., 1994).
A thorough chemical and physical understanding of DNA recognition by transcription factors requires a combination of structural and thermodynamic studies. The precise recognition of a defined DNA sequence by a given transcription factor necessitates optimal shape complementary between the interacting species. The binding reactions are thereby characterised through the formation of large complementary surfaces at the interface between the DNA and the protein. While rigid body associations have been observed, significant conformational changes of both the DNA and the protein often occur on complex formation to ensure an optimal fit. It is therefore important to understand both the intrinsic and the induced properties of the associating molecules.

While most prokaryotic transcriptional regulators bind DNA with high specificity, many eukaryotic transcription factors display only modest DNA binding specificity. The amount of DNA binding specificity of a protein depends on the conformational flexibility of its DNA recognition element. The recognition helices of the eukaryotic bHLH- and bHLHZ- motifs adopt well defined structures only upon binding to DNA. However, even in the complexes, large portions of the recognition helices face the solvent. The conformational flexibility of such recognition elements can be reduced and their DNA binding specificity altered through interaction with other components of the transcriptional machinery. Therefore, transcriptional regulation in eukaryotes relies to a large part on multiprotein complexes with the potential for combinatorial interactions.

A detailed knowledge of the molecular basis of the DNA binding specificities of transcription factors is critical to our understanding of their biological function. In Chapter 2 an approach based on knowledge and rational design is
used with the aim of producing a set of mutants of the protein E12 that show higher specificity on a heterologous sequence of DNA than the wild-type protein.

A different approach to increase specificity is to connect the two subunits through a covalent linker to form a 'single chain dimer', as described in Chapter 3. Earlier work has shown that connecting two MASH-BHLH units covalently through a disulfide bridge at the C-terminal end of helix 2 results in increased affinity and specificity (Künne & Allemann, 1997). 'Single chain multimers' therefore provide an approach to the creation of hybrid proteins with novel properties, such as increased specificity and activity. The advantage of single chain fusions over disulfide bridged dimers is that they can be displayed on the surface of filamentous phages and novel DNA-binding properties can be selected from a large repertoire of mutant proteins.

In Chapter 4 and 5, the binding of MASH-BHLH and MM17 to MCK-S (Buskin & Hauschka, 1989) and to heterologous DNA is studied by isothermal titration calorimetry (ITC). Thermodynamic studies provide insight into the forces that drive the association between proteins and DNA (Ladbury & Chowdhry, 1996). Together with structural studies, they help us to understand the association in physical terms.
Chapter 2: Alteration of the binding specificity of E12

2.1 Introduction

The specificity of DNA recognition by proteins is achieved through the complementarity of the interaction surfaces of protein and DNA. In the complex between E47 and an E-box-containing DNA sequence (Ellenberger et al., 1994), an extensive DNA contact surface of 1052 Å² was observed (Figure 2.1). Direct readout of the DNA sequence, i.e., the contribution to specific recognition from interactions between the functional groups of the protein and the bases of the DNA, is generally assumed to be a major determinant of sequence specificity (Seeman et al., 1976; Rosenberg &

Figure 2.1: Sketch of the DNA complex of the E47 dimer. The basic region and helix 1 and 2 are indicated on one subunit.

The crystal structure of the DNA complex of the BHLH protein MyoD revealed a similar set of contacts (Ma et al., 1994). To determine the importance of individual contacts and their contribution to the overall stability of the complex between El2 and DNA, several mutant proteins were produced, purified and their affinity for various DNA sites determined.

2.2 Rational design of a El2-DNA complex with new specificity

Figure 2.2 shows the network of hydrogen bonds and salt bridges important for specificity in the complex of E47 with an E-box sequence. However, only two amino acid residues were found to contact the functional groups of the bases of the E-box. Glu 345, which is strictly conserved in all BHLH proteins (Figure 2.3), makes two direct contacts to the amino groups of C(1) and A(2) of the E-box (Figure 2.4) and a water-mediated contact to N7 of the guanine at position -1 (Figure 2.2, panel A). The position of the glutamate side chain is fixed through van der Waals contacts between the β- and γ-CH₂ groups of Glu 345 and the methyl group of thymine as well as through a salt bridge to Arg 348, a residue which is in direct contact with the phosphate backbone (Figure 2.5). Replacing Glu 345 with a glutamine would result in the loss of a hydrogen bond either to the amino group of C(1) or that of A(2) (Figure 2.2, panel B). Weaker binding of the protein to MCK-S would be expected. Also, the complex with wild-type El2 and the oligonucleotide SQ, where A(2) is replaced by a guanine, can form one less hydrogen bond (Panel C). A(2) was changed rather than C(1), because G(1), which forms a basepair with C(1),
makes contacts with Asn 341 (Figure 2.6). A similar set of contacts as in the wild-type could be formed in the complex of EQ with the oligonucleotide SQ (Panel D), so that affinity and specificity could be retained.

Figure 2.2: Panel A: contact network according to the crystal structure of E47 with DNA (Ellenberger et al., 1994). Panels B-D show possible contacts in the complexes with the mutant protein EQ and the SQ DNA sequence; EQ/MCK-S (B), E12/SQ (C), EQ/SQ (D).
Figure 2.3: Sequence similarities of MASH, E12 and other BHLH proteins. The alignment is grouped into the neurogenic, the myogenic, and the E12/E47 subfamilies. Amino acids are given in the single letter code, and conserved residues are in boldface.
Figure 2.4: Sequences of the oligonucleotides used in the EMSAs. The E-box sequence is underlined and the mutations relative to MCK-S are indicated in bold face. The numbering system used for the E-box sequence is indicated. P = Purin.

Figure 2.5: Networked hydrogen bonding interactions between Glu 345, Arg 348 and the CpA dinucleotide of the E-box sequence. Glu 345 is hydrogen bonded to N4 of C(1) and to N6 of A(2). Arg 348 forms a 'clamp' which connects the phosphate backbone to Glu 345, thereby locking the conformation of the side chain of Glu 345. The β- and γ- CH₂- groups of Glu 345 make van der Waals contacts to the methyl group of T(2'). The program MacMoMo (Dobler, 1992) was used to create this display from the coordinates of the DNA complex of E47 (Ellenberger et al., 1994).
2.3 Construction of the E12 mutants

Mutations in the E12 gene were introduced through the Kunkel method (Kunkel, 1985; Kunkel et al., 1987) or PCR-based site-directed mutagenesis. For a detailed description see materials and methods.
The gene coding for E12 was cloned into the vector pBSE (modified Bluescript pBSK\(^+\)) (Meierhans, Thesis 1998) resulting in the plasmid pBSE-E12. E.coli CJ236 competent cells were transformed with pBSE-E12. CJ236 cells contain a mutation in the *dut* (dUTPase) and the *ung* (uracil N-glycosylase) genes. Therefore the incorporation of uracil into DNA becomes possible. The cells were grown overnight at 37°C on an agar plate containing ampicillin and chloramphenicol.

Single colonies were picked and grown in liquid media. Adding helper phage M13-K07 produced phagemids with uracil-containing DNA. The phagemid DNA was extracted and used for the mutagenesis reaction. The mutagenic primer E12-E(345)Q (Figure 2.7) was then annealed to the single-stranded phagemid DNA and the second strand was produced using T7 DNA Polymerase and T4 DNA ligase. After transformation into DH5\(\alpha\) cells, the uracil-containing strand was inactivated and the strand with the mutation further replicated. The DNA sequences were verified using the dideoxy sequencing method (Sanger et al., 1977).

For all other mutants, PCR-mediated site-directed mutagenesis was used (Figure 2.8). In the first step, two PCRs were performed. In PCR A, the T7 and 3' mutation primer were used to amplify the gene of E12 between the T7 promotor of the plasmid pJGetita and the site of the mutation. PCR B amplified the gene from the site of mutation to be introduced to a DNA sequence 3' (Get3far) of the plasmid pJGetita (Figure 2.7). The two products were purified by agarose gel electrophoresis. In a second step, the purified DNA fragments were mixed in equimolar amounts and a further PCR was performed without primers. After 3 cycles, T7 and Get3far primers (Figure 2.7) were added and the PCR continued for an additional 26 cycles. The
resulting product was purified, digested with NdeI and BamHI and cloned into the vector pJGetit resulting in the plasmids pJGetE12-E(345)Q, pJGetE12-E(345)L, pJGetE12-R(348)K or pJGetE12-R(348)M.

E12-E(345)Q: 5'-AAC GCC CGA CAG CGC CTG CGT-3'
E12-E(345)L: 5'-AAC GCC CGA TTG CGC CTG CGT-3'
E12-E(345)L.2: 5'-ACG CAG GCG AAAC TCG GGC GTT-3'
E12-R(348)K: 5'-CGA GAG CGC CTG AAA GTC CGC GAC -3'
E12-R(348)K.2: 5'- GTC GCG GAC TTT CAG GCG CTC TCG-3'
E12-R(348)M: 5'-CGA GAG CGC CTG ATG GTC CGC GAC-3'
E12-R(348)M.2: 5'-GTC GCG GAC CAT CAG GCG CTC TCG-3'
T7 start: 5'-TAA TAC GAC TCA CTA TAG GG-3'
Get3far: 5'-AAT GCA GAT CCG GAT ATA GT-3'

**Figure 2.7:** Primers used for mutagenesis. The codons affected are underlined and the mutations in bold face.
2.4 Expression and Purification of E12 and mutants

The plasmid pJGetE12 containing the gene coding for E12 was transformed into BL21(DE3)pLysS cells which were then plated on an agar plate containing ampicillin and chloramphenicol. Colonies were picked and grown
overnight in 2xYT medium at 37°C. 2xYT (0.5 l) medium was inoculated with the overnight culture. Antibiotics were added. The cells were grown at 37°C, with shaking, to an OD$_{600}$ of 0.5, then induced with IPTG. A 1 ml sample was removed before induction and after 90 and 180 minutes, respectively, for gel analysis. Cells were centrifuged, the supernatant was discarded and the pellet stored at -20°C. The samples were analysed on an SDS-PAGE gel. (Figure 2.9)

![E12 Protein Expression](image)

**Figure 2.9:** Expression of the E12 protein. Samples were collected after 0, 90 and 180 min and run on a 12% polyacrylamide gel. The gel was stained with Coomassie Blue. The arrow indicates authentic E12 protein used as a marker.
The cells were thawed in the presence of water and resuspended, then lysis buffer was added. The suspension was sonicated and then centrifuged. The pellet was resuspended in lysis buffer and recentrifuged. This process was repeated once, then the pellet was washed twice with 5 mM sodium acetate (pH 5.0) and centrifuged. The pellet was then dissolved in urea buffer with sonication and the sample was centrifuged. The supernatant was applied to a column containing Bio-Gel CM A ion-exchange resin. The column was washed extensively with urea buffer and the protein eluted with urea buffer containing 1 M sodium chloride. Fractions containing the desired protein were combined and dialysed overnight against urea buffer. The protein was further purified by preparative HPLC on a Resource-S sulfonate ion-exchange column using a linear gradient from urea buffer to urea buffer containing 1 M NaCl in urea as a gradient of 1 %/min.

Figure 2.10: Ion exchange chromatography of E12. The protein was eluted from a Resource-S cation exchange HPLC column with 1 M NaCl in urea as a gradient of 1 %/min.
sodium chloride (Figure 2.10). The collected fractions were pooled and concentrated by ultrafiltration and the buffer was exchanged with 5 mM sodium acetate pH 5.0 and 5 mM DTT by dialysis. The purified protein, that was approximately 97% pure, was stored at -20°C.

2.5 Mass spectrometry of E12 and mutants

The masses of the purified E12 and E12 mutant proteins were confirmed by MALDI-TOF mass spectrometry. The spectra in the Figures 2.11 to 2.14 show the expected masses for the proteins. The mass spectrometer was calibrated with myoglobin as a standard. The spectrum of myoglobin was either taken before or together with the E12 or the mutant proteins.

![MALDI-TOF mass spectrum](image)

**Figure 2.11:** MALDI-TOF mass spectrum of E12 measured together with myoglobin as standard. The singly, doubly, triply and quadruply charged E12 (41166.2, 20575.9, 13725.7, 10303.6) and the singly, doubly and triply charged myoglobin (16941.5, 8470.7, 5650.9) were detected.
Figure 2.12: MALDI-TOF mass spectrum of E12-E345L. The peaks at 41166.8, 20582, 13714.2 and 10295 correspond to the singly, doubly, triply and quadruply charged E12-E345L. Myoglobin was detected at 16946.7 (singly), 8473.1 (doubly) and 5645.4 (triply charged).

Figure 2.13: MALDI-TOF mass spectrum of E12-R348K. At 41147, 20598, 13736 and 10311 the singly, doubly, triply and quadruply charged E12 mutant protein was detected.
2.6 Binding affinities of the E12 mutants

In electrophoretic mobility shift (EMSA) titration experiments (Figure 2.15), the apparent dissociation constants (K_D) of the complexes between E12 and a range of oligonucleotides (Figure 2.4) were measured. The K_D of the complex between E12 and the MCK-S oligonucleotide was 2.7 (±0.55) x 10^{-15} M^2 at
4°C, corresponding to a free energy change of -19.85 kcal/mol for the binding process (Table 1). The DNA binding specificity of E12 was determined to be approximately 100 (Table 1 and Figure 2.16) (Meierhans et al., 1995). The binding specificity was defined as the average of the binding-activities of E12 to SP-1 and NoEBox divided by the affinity of E12 to MCK-S.

![Figure 2.15: Autoradiogram of an electrophoretic mobility shift assay (EMSA) titration of E12 with the MCK-S oligonucleotide. Lane 1: MCK-S alone; lanes 2-8: increasing amounts of E12; [MCK-S] = 10nM.](image)

The dissociation constant of the complex of E12 with the P2 oligonucleotide (Figure 2.4), in which A(2) and A(5) of MCK-S were replaced with purine, which cannot form any hydrogen bonds, was approximately four times higher than that of the MCK-S complex. Removal of the amino group at C6 of
adenine reduced the binding free energy of the E12 complex by approximately 0.9 kcal/mol.

Figure 2.16: Determination of the apparent dissociation constants of the complexes E12, E12-E345L, E12-E345Q, and E12-R348K with the oligonucleotides MCK-S (○), MCK-SQ (△), and NOEbox (□). The fraction θ of DNA bound was determined as the activity of the retarded band divided by the sum of the activities of the retarded and unretarded bands from EMSA titrations.
When A(2) was replaced with guanine (oligonucleotide MCK-SQ) (Figure 2.4), the stability of the complex was diminished further (Table 1 and Figure 2.16), presumably because of a repulsive interaction between O4 of guanine and one of the carboxylate oxygens of glutamate (Figure 2.5). However, replacing Glu 345 with glutamine increased the affinity to the levels of the E12-MCK-S complex (Table 1). Again, when purines were introduced in the 2- and the 5-position of the oligonucleotide, the interaction with E12-E345Q was weakened by approximately 1.1 kcal/mol, indicating that the hydrogen bonds between the carboxylate of Glu 345 and N4 of adenine and between the amide of Gln 345 and O4 of guanine in MCK-SQ contribute approximately the same amount to the stability of the respective complexes.

Another aspect of the results obtained with E12-E345Q was surprising. Not only was the MCK-SQ complex of E12-E345Q more stable than the complex with wild-type E12, but complexes with the heterologous oligonucleotides SP-1 and NOEbox also displayed significantly enhanced stability (Figure 2.16 and Table 1). The concentrations for half-maximal binding of E12-E345Q to MCK-SQ, SP-1, and NOEbox were 38nM, 61 nM, and 56 nM, respectively. The specificity of DNA binding of E12-E345Q was therefore only approximately three (Table 1). However, the DNA binding affinities of E12-E345Q were comparable to the affinity of wild-type E12 for E-box-containing DNA sequences (Table 1). It appears that BHLH-proteins like E12 select E-box containing DNA from heterologous DNA sequences mainly by destabilising the interaction with the latter. Since Arg 348 of E12 connects the phosphate backbone to the E-box through a set of hydrogen bonds via Glu 345, this arginine residue plays a central role in specific DNA binding (Figure
2.2). While a hydrogen bond can still be formed between the amide oxygen of Gln 345 of E12-E345Q and N\textsubscript{348} of Arg 348 in the complex with MCK-SQ, the absence of the negative charge at position 345 most likely allows the arginine side chain to move closer to the phosphate 5' of C(1). This enhanced interaction between the phosphate and Arg 348 is probably the reason for the high DNA binding affinity observed with E12-E345Q.

Table 1: Apparent dissociation constants of the complexes between E12, E12-E345Q, E12-E345L, and E12-R348K and various oligonucleotides determined by EMSA.

<table>
<thead>
<tr>
<th></th>
<th>E12</th>
<th>E12-E345Q</th>
<th>E12-E345L</th>
<th>E12-R348K</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCK-S</td>
<td>2.7(±0.55)</td>
<td>1.04(±0.69)</td>
<td>17.0(±8.36)</td>
<td>628(±85.8)</td>
</tr>
<tr>
<td>SP-1</td>
<td>170.0(±40.0)</td>
<td>2.59(±0.79)</td>
<td>47.3(±35.6)</td>
<td>1437(±556)</td>
</tr>
<tr>
<td>NoEBox</td>
<td>370.0(±50.0)</td>
<td>3.17(±1.26)</td>
<td>30.5(±0.23)</td>
<td>3793(±690)</td>
</tr>
<tr>
<td>MCK-SQ</td>
<td>26.2(±10.0)</td>
<td>1.46(±0.47)</td>
<td>34.2(±1.10)</td>
<td>3115(±138)</td>
</tr>
<tr>
<td>P2</td>
<td>11.8(±7.5)</td>
<td>9.60(±1.90)</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>

nd: not determined
Not surprisingly, replacing Glu 345 with leucine led to a mutant that displayed little specificity (Table 1). On the other hand, the DNA binding affinity was reduced by only approximately one order of magnitude when compared to MCK-S binding to E12, again presumably because of a better interaction between Arg 348 and the phosphate group.

Changing Arg 348 to lysine reduced the affinity of MCK-S binding approximately 230-fold (Table 1). E12-R348K bound to MCK-S half-maximally at a concentration of 792 nM, while 1.12 μM and 1.94 μM E12-R348K were required to bind 50% of the SP-1 or the NOEBox DNA (Table 1). In the absence of the 'clamp' formed by Arg 348, Glu 345 is not kept in the proper position to form optimal contacts to C(1) and A(2) (Figure 2.5). The shorter side chain of lysine prevents it from fulfilling a 'clamp' function. The distance between the N^ of Arg 348 and the oxygen of the phosphodiester is 3.17 Å in the wild-type complex, while 3.02 Å separate the other N^ from the oxygen of the glutamate side chain (Figure 2.5). For all dihedral angles in the side chain of Arg 348 the closest possible distance between the carbon of the guanidinium group (which corresponds to the N_e group of lysine) and the carboxylate is 3.31 Å. In this orientation the distance to the oxygen in the phosphodiester would be increased to 4.4 Å (Figure 2.17). When the guanidinium carbon is brought to its closest possible distance to the phosphate oxygen, the separation from the carboxylate of E(345) is increased to 5 Å.
Figure 2.17: Distances between the carboxylate of Glu 345 and the phosphate oxygen to the Nε of lysine in the R348K mutant, respectively. 3.31Å is the closest distance of Nε of lysine to the carboxylate of Glu 345. This position increases the distance of Nε of lysine to the phosphate to 4.4Å.

Therefore, in neither of these two orientations could the side chain of Lys 348 of E12-R348K perform a clamping function analogous to Arg 348 in the wild-type complex. Glu 345 could not be firmly oriented towards the CpAdinucleotide, resulting in the significant loss of DNA binding specificity observed with E12-R348K. The effect on DNA binding specificity caused by removing the charge altogether upon replacement of Arg 348 with methionine could not be analysed because the DNA binding affinity of E12-R348M was too small to be measured by EMSA (Kd>10^{-11} M^2).
Chapter 3: A Single Chain 'Dimer' of MASH-BHLH²

3.1 Design of a bHLH single chain 'dimer'

The DNA binding specificity displayed by BHLH proteins was found to be small. The BHLH domain of MASH-1 binds to E-box-containing DNA with only marginally higher affinity than to heterologous DNA (Meierhans et al., 1995; Künne et al., 1996). However, covalently linking the subunits of MASH-BHLH through the introduction of a disulfide bond at the C-terminal end of helix 2 increased the DNA binding specificity by one order of magnitude (Künne & Allemann, 1997). The linkage enforced the close proximity of the two helix 2 regions of the individual subunits. In sharp contrast to wild-type MASH-BHLH, the crosslinked 'dimer' was found to be stably folded, even in the absence of DNA.

The subunits of many multimeric protein complexes can be connected through the introduction of covalent linkers. For example, fusion of the α- and β-subunits of glycyl-tRNA synthase via a short peptide linker (Toth & Schimmel, 1986) and fusion of the NS3 protease of the hepatitis C virus with its cofactor peptide NS4A (Dimasi et al., 1998) created fully active single-chain proteins. Other examples include CuZn superoxide dismutase (Hallewell et al., 1989), avian retroviral proteases (Bizub et al., 1991), the RNA binding protein ROP (Predki & Regan, 1995), the sweet tasting peptide monellin (Kim et al., 1989), the human chorionic gonadotropin (Narayan et al., 2000), single chain antibodies (Tang et al., 1996; Huston et al., 1996; Johnson & Bird,

1996, Winthrop et al., 1999; Berger et al., 1999) and both the 434 and arc repressors (Robinson & Sauer, 1996; Percipalle et al., 1995).

Single chain dimers of MASH-BHLH provide an opportunity to address several questions concerning molecular recognition. In particular, creation of such hybrid proteins could permit engineering of novel DNA binding properties. Since amino acids in the two domains can be varied independently, it should be possible through mutagenesis to direct the single chain dimers to asymmetric DNA target sequences. In addition, single chain dimers can be displayed on the surface of filamentous phage particles and new DNA binding properties can be selected for through random mutagenesis from large repertoires of mutant proteins (Tang et al., 1996; McCafferty et al., 1990, Winter et al., 1994; Burton & Barbas, 1994; Jung et al., 1999; Zhang & Davidson, 1999). Random mutagenesis and in vitro selection by phage display has been used to create variants of zinc finger proteins with altered DNA binding properties (Rebar & Pabo, 1994, Jamieson et al., 1994; Fairbrother et al., 1998; Atwell & Wells, 1999). These experiments were greatly facilitated by the monomeric nature of these transcription factors.
3.2 Construction of MM8 and MM17

Single chain dimers of MASH-BHLH or MASH-BHLH with E12-BHLH were constructed by connecting the C-terminus of one subunit with the N-terminus of the second producing a monomer with a linker of varying length.
MM8 and MM17 contain a linker of 8 and 17 amino acids, respectively (Figure 3.1). Glycines were chosen for flexibility and threonine and serine will help to keep the protein soluble. The linkers are shown in Figures 3.5 and 3.6 with the amino acid single letter code in bold face.

The construction strategy is illustrated in Figure 3.2.

Plasmid pJGetitMASH-BHLH, which contains a fragment of the MASH-1 cDNA coding for the BHLH domain from G(106) to D(172) (Meierhans et al., 1995), was digested with the restriction enzymes PstI and BamHI. The resulting vector fragment was ligated with a cassette (bHLHlink1) resulting in plasmid pJGetMAlink1. This plasmid contains the first half of the linker and two unique restriction sites for AgeI and KpnI.

BHLHlink1:

<table>
<thead>
<tr>
<th>PstI</th>
<th>AgeI</th>
<th>KpnI</th>
<th>BamHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>g cag ctg ctg ACC GGT GGT ACC GGg</td>
<td>ac gtc gtc gac gac TGG CCA CCA TGG CCC cta g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QQLLTGGTG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lower case letters indicate bases from the coding region of the MASH-1 gene. Capitals correspond to bases of the introduced part of the linker.

In a second step, the cassette bHLHlink2, coding for the second half of the linker, was inserted into the NdeI site of pJGetMASH-BHLH to give plasmid pJGetMAlink2. The AseI site was destroyed by insertion.
BHLHlink2:

\[
\begin{array}{cccccccc}
NdeI & \text{KpnI} & \text{AseI} \\
T & ATG & GGT & ACC & GGG & GGT & GGA & AGT & AT \\
AC & CCA & TGG & CCC & CCA & CCT & TCA & TAA & T \\
M & G & T & G & G & G & S & I \\
\end{array}
\]

Finally, the KpnI-BamHI fragment of the insert in pJGetMAblink2 was inserted between the KpnI and BamHI sites of pJGetMAblink1 to yield pJGetMM8.

To construct the plasmid with the gene that encodes the protein MM17, pJGetMM8 was digested with the restriction enzymes AgeI and KpnI. The DNA sequence coding for the linker extension was constructed by ligating the vector fragment with the double-stranded oligonucleotide ZS1.

ZS1:

\[
\begin{array}{cccccccc}
\text{AgeI} & \text{KpnI} \\
CC & GGT & GGA & GGT & AGT & GGT & GGC & GGG & TCA & GGT & GGA & GGT & AC \\
A & CCT & CCA & TCA & CCA & CCG & CCC & AGT & CCA & CCT & C \\
G & G & G & S & G & G & S & G & G & G \\
\end{array}
\]

The DNA sequence of all constructs was verified using the dideoxy sequencing method (Sanger et al., 1977). The DNA and amino acid sequences of MASH-BHLH and MM17 are displayed in Figures 3.3 and 3.5.
Figure 3.2: Construction of the single chain dimers
3.3 Construction of ME8 and ME17

Plasmid pJGetME8 was constructed analogous to pJGetMM8, except in the second step where pJGetE12-BHLH was digested with NdeI.
To construct plasmid pJGetME17, the Agel-KpnI fragment of pJGetME8 was ligated with ZS1, the DNA sequence coding for the linker extension (Figure 3.3). The DNA and amino acid sequences important for the cloning of ME17 are displayed in Figures 3.3 - 3.6.

Figure 3.3: DNA sequence of MASH-BHLH and single letter code for the translated gene. Restriction sites important for the construction of MM17 are underlined and indicated.

```
NdeI
| CATATGGGCTACGCTTCCACAGCAGCAGCCGGCAGCCGTCCAGCGCGCCGCAACGAGGCG
-------------------------+-----------------------------
GTATACCCGATGTCGGGAAGGTGTGTCGTCGGCCGGTCGCGCGACCGGGCGCAGCGTGGGCG
MGYSLPQQQPAAVARRNER

GAGCGCAACCGGTCAAGTTGTTAACCTGGCCTTGGCACCCTCCGGGAGCATGTCCCC
61 +-------------------------------+
CTCACCACGTGGCCAGTTCAACCAATTGGGACGAAACGGAGCTGGAGAAGGCCCCGCTGACAGGGG
ERNRVRVKLVNLGFAATLREHV

AACGCCGCTGCAACAAAGAGATGAGGAGACGTGGCTGCCGGGATCGAATAC
121 +-------------------------------+
TTGCGTGAGGCTTTCTCTCTCTCTGTTACCTCGCAGCGAACAGCCGACCGCTGATATG
NGAANKKMSKVEETLRSAY
```
Figure 3.4: DNA sequence of E12-BHLH and corresponding translation to amino acid single letter code.
Figure 3.5: DNA sequence of MM17 and amino acid single letter code.

**BamHI**

GCCGTGGCCGTATCCTCAGGCTGGACAGCTGGCCGATGAGGATCC
CGGCACCCGCAAGCGGATCGACCTCGTCGTCGCTCAGCTACCTCTAGG

<table>
<thead>
<tr>
<th>181</th>
<th>228</th>
</tr>
</thead>
<tbody>
<tr>
<td>A V A V I L S L E Q Q V R * G S -</td>
<td></td>
</tr>
</tbody>
</table>

**NdeI**

CATATGGGCCTACGCTCCACAGCAGCGCCAGCGGTGGCCGACGCAGGC
GTATACCCGATGTCGGAAGGTGTCGTCGTCGGCCGTCGGCACCGCGCGGCGTTGCTCGCG

<table>
<thead>
<tr>
<th>1</th>
<th>61</th>
</tr>
</thead>
<tbody>
<tr>
<td>M G Y S L P Q Q Q P A A V A R R N E R -</td>
<td></td>
</tr>
</tbody>
</table>

**AgeI**

ATCGCGCGCTGACGAGCGTCTGGGAGTGATGTGTGGCCGAGCTGGGTAGTGACCTGGCCGCTAG

<table>
<thead>
<tr>
<th>121</th>
<th>181</th>
</tr>
</thead>
<tbody>
<tr>
<td>N G A A N K K M S K V E T L R S A V E Y -</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>181</th>
<th>228</th>
</tr>
</thead>
<tbody>
<tr>
<td>I R A L Q Q L L T G G G S G G S G G G G -</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.6: DNA sequence of ME17 and amino acid single letter code.

**KpnI**

ACCGGGGGTGGAAGTATTATGGGCTACAGCCTTCCACAGCAGCAGCCGGCAGCCGTGGCG

TGGCCCCCACCTCTCAATAATACCCGATGTCGGAAGGTGTCGTCGTCGGCCGTCGGCACCGC

<table>
<thead>
<tr>
<th>T G G G S</th>
<th>I M G Y S L P Q Q Q P A A V A</th>
</tr>
</thead>
</table>

301

CGCCGCAACGAGCGCGAGCGCAACCGGGTCAAGTTGGTTAACCTGGGCTTTGCCACCCTC

GCGGCGTTGCTCGCCGCTCGCTCGTTGGCCCAGTTCAACCAATTGGGACCCGAAACGGTGGGAG

| R R N E R E R N R V K L V N L G F A T L |

361

CGGGAGCATGTCCCCAACGGCGCTGCCAAACAAGAAGATGAGCAAGGTGGGAGACGCTGGT

GCCCTCTGACAGGGGTGCCGCGCAGGTTTCTTTCTACTCTGGTCCACCCTCTCCGACGCA

| R E H V P N G A A N K K M S K V E T L R |

421

TCGGCGGTCGAATAATCATCCCGCGGCTGCAGCAGCTGCTGGATCCC

AGCCGCCAGCTTATGTAGGGCCGCCGACGTCTCGACGACCTAGGG

| PstI BamHI |

| TCGGCGGTCGAATAATCATCCCGCGGCTGCAGCAGCTGCTGGATCCC |

465

AGCCGCCAGCTTATGTAGGGCCGCCGACGTCTCGACGACCTAGGG

| S A V E Y I R A L Q Q L L D P |

**NdeI**

CATATGGGCTACAGCCTTCCACAGCAGCAGCGCAGCCGGCAGCCGGCAGCCGGCAGCCG

GTATACCCGATGTCGGAAGGTGTCGTCGTCGGCCGTCGGCACCGC

| M G Y S L P Q Q Q P A A V A R R N E R |

49
3.4 Expression of MM17

Plasmid pJGetMM17 containing the gene coding for MM17 was transformed into BL21(DE3)pLysS cells and plated on agar containing ampicillin and chloramphenicol. Colonies were picked and grown at 37°C. 2xYT medium (0.5 l) were inoculated with overnight culture. The cells were grown at 37°C with shaking to an OD_{600} of 0.5, then induced with IPTG. Samples (1 ml) were taken before induction and after 60, 120 and 180 minutes for gel analysis. Cells were centrifuged. The supernatant was discarded and the pellet stored at -20°C. Every 1 ml sample taken for gel analysis was centrifuged and the supernatants were discarded. The samples were vortexed and sonicated. After heating to 95°C, the samples were loaded and run on an SDS-PAGE gel. (Figure 3.7)
Figure 3.7: Expression of MM17. Lanes A-D show the expression of MM17 (16.5 kDa) at \( t = 0, 60, 120 \) and 180 min. Lane M shows the molecular weight marker (Sigma).

### 3.5 Purification of MM17

Cells were thawed in the presence of water, then resuspended by vortexing or thorough shaking. Lysis buffer was added. The suspension was sonicated at room temperature and then centrifuged. The pellet was resuspended in lysis buffer and recentrifuged. After this process was repeated once, the pellet was washed twice with 5 mM sodium acetate (pH 5.0) and centrifuged. The pellet was dissolved in urea buffer. To help dissolving, the sample was sonicated, then centrifuged. The solution was applied to a column containing Bio-Gel CM A ion-exchange resin. The column was washed extensively with urea...
buffer and the protein eluted with urea buffer containing 1 M sodium chloride. Fractions containing the desired protein were combined and dialysed against buffer A overnight. The protein was further purified by preparative HPLC on a Resource-S sulfonate ion-exchange column using a linear gradient from urea buffer to urea buffer containing 1 M sodium chloride (Figure 3.8).

**Figure 3.8:** Ion exchange chromatography of MM17. The protein was eluted from a Resource-S cation exchange HPLC column with 1 M NaCl in urea as a gradient of 1%/min.
The collected fractions were pooled, concentrated by ultrafiltration, and the buffer was exchanged with 5 mM sodium acetate (pH 5.0) containing 5 mM DTT. The protein was stored at -20°C. Figure 3.9 shows MM17 protein obtained in crude extracts and after purification.

Figure 3.9: SDS-PAGE of crude extracts of E. coli cells harbouring an expression plasmid for MM17 just before (lane A) and 2 hours after induction of expression (lane B). The arrow indicates overproduced MM17. Lanes C to E show the purified proteins MM8, MM17 and MASH-BHLH. Mobilities of molecular weight marks (MW) are given in kDa.
Figures 3.10 - 3.12 show the MALDI-TOF mass spectra of MM8, MM17 and ME17, respectively. The molecular masses corresponded well with the calculated masses of MM8, MM17 and ME17, respectively, without the N-terminal methionine. Edman degradation gave the correct N-terminal sequences and confirmed that the N-terminal methionine had been removed proteolytically.

**Figure 3.10:** MALDI-TOF mass spectrum of MM8. The masses for singly (16001) and the doubly charged (8000.8) protein were detected.

**Figure 3.11:** MALDI-TOF mass spectrum of purified MM17. Calculated mass is 16559.8 without the N-terminal methionine. The peaks at 16580.7 and 8287.9 correspond to singly and doubly charged MM17.
Figure 3.12: MALDI-TOF mass spectrum of purified ME17. The calculated mass is 15922.3 without the N-terminal methionine. The masses of 15926 and 7966.8 correspond to the singly and doubly charged protein.

3.6 CD spectroscopy of MM8 and MM17

CD spectroscopy was used to obtain structural information about MM8 and MM17. Compared to calculated circular dichroism containing varying percentages of alpha-helix and random coil (Greenfield & Fasman, 1969; Comton & Johnson, 1986), the CD spectrum of a 1 μM solution of wild-type MASH-BHLH revealed that approximately 35% of the amino acids were in an alpha-helical conformation (Figure 3.13). The single chain dimers showed an alpha-helical content of about 45% and, even though the content of alpha-
helicity was higher, a significant portion of the peptides remained unstructured. This is in contrast to the behaviour of the MASH mutant MASH-GGC, in which the BHLH subunits are held together through a disulfide bond at the C-terminal end of helix 2 under oxidizing conditions (Künne & Allemann, 1997). MASH-GGC was stably folded and mainly alpha-helical. The disulfide linkage keeps two segments of the BHLH domain in close proximity, which in the folded "dimer" are in direct contact. In the

![Figure 3.13: Circular dichroism spectra of MASH-BHLH, MM17 and MM8 and of their complexes with MCK-S. Concentrations were 1 μM MASH-BHLH and 0.5 μM MM8 or MM17 (corresponding to 1 μM of BHLH equivalents). MCK-S was 0.5 μM.](image)

- a) MCK-S only
- b) MASH-BHLH
- c) MM17
- d) MM8
- e) MM8 and MCK-S
- f) MASH-BHLH and MCK-S
- g) MM17 and MCK-S

According to the figure:

- a) MCK-S only
- b) MASH-BHLH
- c) MM17
- d) MM8
- e) MM8 and MCK-S
- f) MASH-BHLH and MCK-S
- g) MM17 and MCK-S
"single chain dimers", two parts of the protein are held together that are remote from each other even in the folded conformation.

MASH-BHLH undergoes a concentration dependent transition from a mainly unfolded monomer to a stably folded dimeric form with a dimerisation constant of about 2 μM (Künne & Allemann, 1997). The CD spectra of MM8 and MM17 were essentially unchanged over the concentration range 0.1-5 μM, which corresponds to 0.2-10 μM monomer equivalents. This is expected for a unimolecular folding reaction (Figure 3.14). No evidence for significant "crossfolding" of the BHLH subunits of MM8 and MM17 to form dimeric species or higher aggregates was observed.

![Figure 3.14: Molar ellipticity at 222 nm, [\(\Theta\)]_{222}, from CD spectra of MM17 as a function of the concentration of MM17.](image-url)
Upon addition of 1 equivalent double-stranded oligonucleotide containing an E box sequence to a solution of MM8 or MM17, a folding transition from a largely unfolded to a more alpha-helical conformation was observed (Figures 3.13 and 3.15). A similar change in the CD spectrum occurred when MCK-S was added to wild-type MASH-BHLH (Figure 3.13; Künne & Allemann, 1997; Meierhans et al., 1995). Interestingly, the degree of helicity observed in the different complexes varied. In the DNA complex of MM17, 78% of all residues were in an alpha-helical conformation, an increase of 5% when compared with the wild-type complex (Figure 3.13). On the other hand, the percentage of alpha-helicity was ~75% in the MM8 complex (Figure 3.13). This might be a consequence of the shorter length of the linker used in MM8. Either the N-terminal end of the basic region or the C-terminal part of helix 2 might have to unfold partly to allow proper folding of MM8 on the DNA. However, if so, this local unfolding did not diminish the DNA binding affinity of MM8 (Table 2).

The structural changes upon DNA binding observed in both wild-type MASH-BHLH and the "single chain dimers" were in sharp contrast to the behaviour of disulfide-linked MASH-BHLH, which was fully folded even in the absence of DNA. No conformational change was observed when DNA was added (Künne & Allemann, 1997), indicating that the processes of dimerization, folding and DNA binding were uncoupled. MM8 and MM17, on the other hand, behave similarly to wild-type MASH-BHLH, in that folding and DNA binding remain coupled processes. Since the two subunits are covalently linked in the single chain dimers, no dimerization occurs on DNA binding. However, the subunits of MM8 and MM17 still undergo a conformational
rearrangement which brings the two subunits into the intimate contact needed for formation of the proper complex.

**Table 2:** DNA binding parameters of MASH-BHLH, MM8 and MM17 measured by EMSA

<table>
<thead>
<tr>
<th></th>
<th>[P1/2]a (nM)</th>
<th>Kd (10^15)b</th>
<th>ΔGobs c (kcal/mol)</th>
<th>ΔΔGobs d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCK-S</td>
<td>SP-1</td>
<td>MCK-S</td>
<td>SP-1</td>
</tr>
<tr>
<td>MASH-BHLH</td>
<td>458.0 (±91)</td>
<td>520.0 (±129)</td>
<td>209.8</td>
<td>270.4</td>
</tr>
<tr>
<td>MM8</td>
<td>16.2 (±5.5)</td>
<td>44.5 (±1.1)</td>
<td>1.1</td>
<td>7.9</td>
</tr>
<tr>
<td>MM17</td>
<td>22.3 (±6.6)</td>
<td>59.3 (±1.7)</td>
<td>2.0</td>
<td>14.1</td>
</tr>
</tbody>
</table>

a Concentrations of protein for which 50% of the DNA binding sites are filled. Standard deviations from multiple measurements under identical conditions are given in parentheses.

b Dissociation constants are reported relative to monomer equivalents: Kd=[MASH-BHLH]^2 for MASH-BHLH; Kd=(2x [MM8])^2 and Kd=(2x [MM17])^2 for the 'single chain dimers'.

c Reaction free energy for the binding reaction: ΔGobs = -RT ln[MASH-BHLH]^2 for MASH-BHLH; ΔGobs = -RT ln[P]_1/2 for MM8 and MM17. Values are for 20°C.

d ΔΔGobs = ΔGobs(MCK-S) - ΔGobs(SP-1) corresponds to the free energy of transferring a protein molecule from an SP-1 site to an MCK-S site.
Figure 3.15: CD spectra of MM8 in the presence and absence of the oligonucleotides MCK-S and SPI.

a) MCK-S alone   c) MM8 and 1 eq. SP-1
b) MM8 alone     d) MM8 and 1 eq. MCK-S

As observed with MASH-BHLH, the coil to α-helix transition was not only induced through addition of E-box-containing DNA, but also by completely heterologous DNA (Figure 3.15; Künne et al., 1996; Künne & Allemann, 1997; Meierhans et al., 1995; Künne et al., 1998). Interestingly, the complex of MM8 with MCK-S showed slightly more alpha-helicity than the complex with heterologous DNA. The same observation was made for the DNA complexes of MM17. While these observations were difficult to interpret, they nevertheless suggested a small difference in the geometry of the specific and
the non-specific complexes of MM8 and MM17. It is noteworthy that no difference in the CD spectra of the specific and non-specific complexes of wild-type (Meierhans et al., 1995) and disulfide-linked MASH-BHLH were observed (Künne & Allemann, 1997).

3.7 DNA binding affinity of MM8 and MM17

Earlier work had shown that MASH-BHLH binds to DNA with moderate affinity and low DNA sequence specificity (Table 2, Meierhans et al., 1995; Künne & Allemann, 1997; Künne et al., 1998). In EMSA titration experiments, apparent dissociation constants were measured for complexes of the single chain dimers with oligonucleotides containing an E-box and with completely heterologous DNA (Figure 2.4). Increasing amounts of the proteins were added to a constant amount of DNA, and the extent of complex formation was measured (Figure 3.16).
The protein concentration at which half of the DNA binding sites are occupied, \([P]_{1/2}\), was determined from the graphs describing the dependence of \(\Phi\), the fraction of DNA bound, on the concentration of the unbound protein (Figure 3.17).
The fraction $\Phi$ 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. Plotting $\Phi$ against the logarithm of the concentration of unbound protein allowed the determination of the concentration $[P]_{1/2}$ at which half of the protein binding sites were occupied (Meierhans et al., 1995). The best fit for DNA binding of "single chain dimers" to the binding isotherm in equation 1 (Langmuir function) was obtained for $n=1$ under the assumption of cooperative binding of a monomer to the DNA (Clore et al., 1982).

![Figure 3.17: Binding isotherms for the binding of MM8 (left) and MM17 (right) to MCK-S (●) and SP-1 (▲) oligonucleotides.](image)

**Equation 1:**

$$\Phi = \frac{1}{(1 + [P]^{1/2}) / ([P]^n)}$$
The single chain dimers bind the MCK-S oligonucleotide half maximally at significantly lower concentrations than MASH-BHLH. While a concentration of 458 (±91) nM was required to occupy 50% of all E-boxes of MCK-S, concentrations of only 16.2 (±5.5) and 22.3 (±6.6) nM were needed for MM8 and MM17 respectively (Table 2). Even when the change from a dimeric to a monomeric species was taken into account, linking of the two BHLH domains lowered the half maximal binding concentration by more than one order of magnitude. The oxidised form of MASH-GGC bound to MCK-S only about 3 times tighter than wild-type MASH-BHLH (Künne & Allemann, 1997).

The ~10 fold increase in the affinities of the single chain dimers for E-box containing DNA sequences compared with wild-type MASH-BHLH could have several different origins. The energies required for stabilization of the single chain dimers could result from additional contacts between the DNA and residues in the protein linker. While this explanation cannot be ruled out based on the existing data, the X-ray structures of the DNA complexes of E47 and MyoD (Ellenberger et al., 1994; Ma et al., 1994) suggest that the residues of the linker pass around one side of the BHLH dimer and that they are shielded from the DNA by one of the BHLH domains (Figure 3.1). An alternative explanation is suggested by isothermal titration calorimetry (ITC) experiments of the DNA binding process of MASH-BHLH (Chapter 4; Künne et al., 1998). While the amino acid residues of the basic region were unfolded in the free proteins even at the high protein concentrations used for ITC (where MASH-BHLH exists mainly as a stable dimer with dimerization mediated through the HLH region), the basic region adopted an α-helical conformation in the complex. In addition, NMR spectroscopy of the BHLH
domain of E47 revealed that the basic region is structurally disordered in the absence of DNA (Fairman et al., 1997). Therefore, the linker which connects the C-terminal end of one BHLH subunit of MM8 or MM17 to the basic region of the other might help position residues in the basic region, thereby improving contacts with the DNA. The linker might reduce the entropic penalty that accompanies folding during DNA binding simply by restricting the conformational freedom of the basic region in the disordered state (Spolar & Record, 1994).

3.8 Specificity of DNA binding

Further evidence that the linker might restrict the conformational mobility of the adjacent basic region was provided by the observation that not only DNA binding affinity but also DNA binding specificity was increased in the single chain dimer when compared with MASH-BHLH. While the affinity for E-box-containing DNA was increased in MM8 and MM17 by 10- to 14-fold, the affinity for heterologous DNA was only 4- to 6-fold higher (Table 2). As a consequence, the free energy of transferring a protein molecule from the heterologous SP-1 DNA to an oligonucleotide containing an E-box was decreased from -0.14(±0.18) kcal/mol for wild-type MASH-BHLH to -0.59(±0.16) kcal/mol for MM8 and to -0.57(0.14) kcal/mol for MM17. Limiting the number of accessible conformations of the basic region through the introduction of the linker could stabilise the complex with specific DNA to a greater extent than the complex with heterologous DNA. Interestingly, while the association of single chain dimers and MCK-S was more exergonic by
~1.2 kcal/mol than the binding reaction of disulfide-linked MASH-BHLH (Table 1, Künne & Allemann, 1997), the specificity increase was slightly smaller. $\Delta\Delta G_{\text{obs}}$ for MM8 and MM17 were -0.59 and -0.57 kcal/mol respectively, while for disulfide-linked MASH-BHLH $\Delta\Delta G_{\text{obs}}$ was -0.71 kcal/mol which is the same within the error of the measurement (Künne & Allemann, 1997).
Chapter 4: Thermodynamics of MASH BHLH/DNA complex formation

4.1 Introduction

Despite the important role played by BHLH proteins in transcriptional activation and cellular differentiation and the many studies of their DNA binding, no comprehensive characterisation of the thermodynamics of their association with DNA exists. Thermodynamic studies provide a description of the forces which drive the reaction between proteins and DNA (Ladbury & Chowdhry, 1996). Together with structural studies, they enable us to understand the association reaction in physical terms. Isothermal titration calorimetry (ITC) was used to study the binding of MASH-BHLH to an E-Box containing oligonucleotide, MCK-S, and to heterologous DNA (Figure 2.4).

4.2 pH dependence of the binding reaction

To monitor the effect of pH on the stability of the DNA complexes of MASH-BHLH with oligonucleotides containing an E-box and with heterologous DNA (Figure 2.4), the dissociation constants of the complexes were measured in electrophoretic mobility shift assays over the pH range from 6.0 to 8.0. These

---

experiments showed that the position of the equilibrium of the binding reaction was only slightly dependent on the pH of the solution under these experimental conditions (Figure 4.1). The dissociation constants for the complexes of MASH-BHLH with both MCK-S and SP-1 DNA were slightly smaller at higher pH, while the specificity of the binding reaction was not dependent on the pH of the solution. These results agreed with the data obtained by isothermal titration calorimetry, which showed that the protonation state did not change on complex formation (vide infra) (Table 3). The majority of the ITC experiments described below were therefore performed at pH 6, a value at which the solubility of MASH-BHLH is sufficiently high even under high salt conditions.

**Figure 4.1:** pH dependence of the interaction between MASH-BHLH and the MCK-S (□) and SP-1 (◆) oligonucleotides measured by EMSA. The concentration of MASH-BHLH, [MASH-BHLH]_{1/2}, at which half of the DNA binding sites are filled is given for the pH range from 6 to 8.
Table 3: Dependence of the Thermodynamic Parameters of the DNA Binding Reaction of MASH-BHLH on Buffer and pH

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH</th>
<th>T (K)</th>
<th>ΔG (kcal/mol)</th>
<th>ΔH (kcal/mol)</th>
<th>TΔS (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MCK-S</td>
<td>SP-1</td>
<td></td>
</tr>
<tr>
<td>MES</td>
<td>6.0</td>
<td>295</td>
<td>-11.1(±0.2)</td>
<td>-10.4(±0.5)</td>
<td>-11.1(±0.2)</td>
</tr>
<tr>
<td>PIPES</td>
<td>6.0</td>
<td>295</td>
<td>-10.7(±0.3)</td>
<td>-9.8(±0.1)</td>
<td>-9.8(±0.1)</td>
</tr>
<tr>
<td>PIPES</td>
<td>7.0</td>
<td>295</td>
<td>-10.6(±0.3)</td>
<td>-10.4(±0.3)</td>
<td>-10.4(±0.3)</td>
</tr>
</tbody>
</table>

\(^a\) Standard deviations from multiple measurements are given.

\(^b\) From the measurement of K_A: ΔG = -RT ln K_A.

\(^c\) Calculated from TΔS = ΔH - ΔG.

4.3 ITC measurements of the binding reaction

Isothermal titration experiments (Wiseman et al., 1989) were used to determine the thermodynamic parameters and the stoichiometry of the binding reaction between MASH-BHLH and the MCK-S and SP-1 oligonucleotides by direct measurement of the evolved heat as a function of temperature in the range from 12°C to 32°C (Table 4, Figure 4.2). For all titrations, the best fit of the experimental data to the binding isotherm was obtained under the assumption of one dimeric MASH-BHLH binding per double-stranded
oligonucleotide (Figure 4.2). Control titrations of MASH-BHLH in buffer which did not contain DNA gave rise to small endothermic peaks of equal size within the limits of error. Titration with a large excess of MASH-BHLH did not reveal any sign of a second binding site, as has been described for other DNA-protein interactions (Berger et al., 1996; Ladbury et al., 1994).

\[ \text{Figure 4.2:} \quad \text{Examples of isothermal titrations of MCK-S (■) and SP-1 (○) with MASH-BHLH. 5 µM DNA was titrated with 184 µM MASH-BHLH in low-salt MES buffer at 290 K.} \]

Abscissa: moles of MASH-BHLH dimer added per mole of DNA target. Control titration of MASH-BHLH into MES buffer (▲).

Inset: raw data for the titration of MCK-S DNA (lower trace); control titration of MASH-BHLH into low-salt MES buffer (upper trace).
Table 4: Thermodynamic Parameters of the Binding Reaction of MASH-BHLH and Different DNA Binding Sites

<table>
<thead>
<tr>
<th>T (K)</th>
<th>$K_A \times 10^{-7}$ (M$^{-1}$)</th>
<th>$\Delta G$ (kcal/mol)$^b$</th>
<th>$\Delta H$ (kcal/mol)</th>
<th>$T\Delta S$ (kcal/mol)$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>285</td>
<td>3.41(±0.1)</td>
<td>-9.8(±0.1)</td>
<td>-19.8(±0.1)</td>
<td>-10.0(±0.2)</td>
</tr>
<tr>
<td>290</td>
<td>4.25(±0.1)</td>
<td>-10.1(±0.2)</td>
<td>-19.9(±1.2)</td>
<td>-9.8(±1.4)</td>
</tr>
<tr>
<td>295</td>
<td>17.75(±5.8)</td>
<td>-11.1(±0.2)</td>
<td>-26.3(±1.8)</td>
<td>-15.2(±2.0)</td>
</tr>
<tr>
<td>300</td>
<td>51.25(±5.1)</td>
<td>-11.9(±0.4)</td>
<td>-29.8(±1.1)</td>
<td>-17.9(±1.5)</td>
</tr>
<tr>
<td>305</td>
<td>41.42(±6.2)</td>
<td>-12.0(±0.5)</td>
<td>-33.2(±1.2)</td>
<td>-21.1(±1.7)</td>
</tr>
</tbody>
</table>

**MCK-S in low-salt MES Buffer (pH 6.0)**

| 285   | 2.08(±0.31)                   | -9.5(±0.1)      | -15.3(±0.1)     | -5.8(±0.2)      |
| 290   | 2.06(±0.22)                   | -9.6(±0.1)      | -17.5(±0.1)     | -7.9(±0.2)      |
| 295   | 2.31(±0.47)                   | -9.8(±0.1)      | -15.5(±1.9)     | -5.7(±2.0)      |
| 300   | 5.14(±6.28)                   | -10.4(±0.5)     | -18.1(±2.5)     | -7.7(±3.0)      |
| 305   | 5.14(±1.93)                   | -10.4(±0.3)     | -25.0(±3.1)     | -14.7(±3.4)     |

**SP-1 in low-salt MES Buffer (pH 6.0)**

| 285   | 2.89(±1.86)                   | -9.9(±0.4)      | -18.4(±0.7)     | -8.5(±1.1)      |
| 295   | 2.55(±1.54)                   | -10.0(±0.5)     | -21.4(±2.9)     | -11.4(±3.3)     |
| 300   | 1.30(±0.02)                   | -9.8(±0.1)      | -22.2(±0.3)     | -12.4(±0.3)     |

**MCK-S in high-salt MES Buffer (pH 6.0)**

| 290   | 3.92(±0.25)                   | -10.1(±0.1)     | -22.0(±0.4)     | -11.9(±0.5)     |
| 295   | 2.85(±0.77)                   | -10.1(±0.2)     | -24.2(±1.6)     | -14.2(±1.8)     |
| 300   | 2.24(±0.78)                   | -10.1(±0.2)     | -29.4(±1.7)     | -19.3(±1.9)     |

**SP-1 in high-salt MES Buffer (pH 6.0)**

| 290   | 2.89(±1.86)                   | -9.9(±0.4)      | -18.4(±0.7)     | -8.5(±1.1)      |
| 295   | 2.55(±1.54)                   | -10.0(±0.5)     | -21.4(±2.9)     | -11.4(±3.3)     |
| 300   | 1.30(±0.02)                   | -9.8(±0.1)      | -22.2(±0.3)     | -12.4(±0.3)     |

$^a$ Standard deviations from multiple measurements are given.

$^b$ From $K_A$: $\Delta G = -RT \ln K_A$.

$^c$ Calculated from $T\Delta S = \Delta H - \Delta G$. 
In low-salt MES buffer, the association constants for the binding reaction between MASH-BHLH and MCK-S and SP-1 were $3.41 (±0.1) \times 10^7 \text{ M}^{-1}$ and $2.08 (±0.31) \times 10^7 \text{ M}^{-1}$ at $12°C$ and $51.25 (±5.12) \times 10^7 \text{ M}^{-1}$ and $5.14 (±1.93) \times 10^7 \text{ M}^{-1}$ at $27°C$, respectively (Table 4). At $27°C$, the concentration of MASH-BHLH needed to bind half of the MCK-S oligonucleotides is therefore approximately 1 order of magnitude lower than the concentration required for half maximal binding of the SP-1 oligonucleotide.

The specificity of the binding reaction of MASH-BHLH was similar in low-salt PIPES buffer (Table 3) and with other heterologous DNA sequences. The specificity of DNA binding was slightly lower in high-salt buffer at pH 6 and in low-salt buffer at pH 7 (Tables 3 and 4). The measurement of the thermodynamic parameters in high-salt buffer at pH 7 was precluded by the low solubility of MASH-BHLH under these conditions.

The melting curves of the MCK-S and the SP-1 oligonucleotide at a concentration of 5 μM, which corresponds to the concentration used in the ITC experiments, were measured by UV spectroscopy and showed a sharp transition at approximately $61°C$ and $68°C$, respectively (Figure 4.3) (Meierhans, Thesis 1998). The hyperchroism was less than 1% at temperatures below $35°C$, indicating that the stability of the duplex DNA did not interfere with the accuracy of the ITC experiments.
Figure 4.3: Thermal stability of the MCK-S (○) and the SP-1 (●) oligonucleotides. The optical melting profile of a 5 μM solution of the double-stranded oligonucleotides was measured at 260 nm. The hyperchromicity is given as a function of the temperature. Hyperchromicity was calculated as $100 \times \left\{ \frac{A_{260}(T) - A_{260}(0^\circ C)}{A_{260}(0^\circ C)} \right\}$.

The results described above were in good agreement with earlier EMSA experiments which had also revealed that MASH-1 binds to DNA with low sequence specificity (Meierhans et al., 1995, Künne & Allemann, 1997). They confirmed in a thermodynamically rigorous assay the earlier finding that MASH-1, and most likely other BHLH proteins, display only modest DNA binding specificity.
It is interesting to discuss the low DNA binding specificity in the context of the crystal structures of the DNA complexes of E47 (Ellenberger et al., 1994) and MyoD (Ma et al., 1994). The high affinity of these BHLH proteins for DNA is a consequence of a large number of contacts formed between the phosphate backbone and amino acid residues. No specific contacts exist between the proteins and the central two base pairs of the E-box (CANNTG). Both in E47 and in MyoD, the carboxylate of a glutamate side chain is hydrogen bonded to the N(4) of cytosine and the N(6) of adenine (Figure 4.4). In both subunits of MyoD, additional hydrogen bonds exist between O(6) and N(7) of the of guanine and one of the N_η of an arginine residue, while in E47 a hydrogen bond is formed between O(6) of guanine and the carbonyl of an asparagine side chain in only one subunit. Obviously, MyoD and E47 rely on only a small number of hydrogen bonds between amino acid residues and the nucleobases for specific DNA recognition, and this might explain the small extent of sequence preference observed with these proteins, at least within the context of the moderate resolution of the X-ray structures.
Figure 4.4: Specific contacts from the Glu 118 of MyoD to N4 of cytosine and N6 of adenine. Additional, a water-mediated contact to N7 of guanine is observed. Arg 121 forms a salt bridge with the carboxylate of Glu 118.

4.4 Calculation of the thermodynamic parameters

Isothermal titration calorimetry experiments allowed the direct measurement of the enthalpy change $\Delta H$ and the association constant $K_A$. Equation 2 relates
K_A to the free energy of binding, $\Delta G$. $\Delta T S$ is obtained from the measured $\Delta H$ and the calculated $\Delta G$ (Equation 3) (Table 4).

The heat capacity change $\Delta C_p$ could be determined from the slope of the linear regression function obtained from the plots of the measured $\Delta H$ as a function of temperature. (Equation 4) (Figure 4.5).

**Equation 2:** $\Delta G = -RT \ln K_A$

**Equation 3:** $T\Delta S = \Delta H - \Delta G$

**Equation 4:** $\Delta C_p = \left( \frac{\partial H}{\partial T} \right)_p$

Spolar & Record have proposed a dissection of the observed entropy change for the association reaction into the entropy change from the hydrophobic effect ($\Delta S_{HE}$), the polyelectrolyte effect ($\Delta S_{PE}$), rotational and translational contributions ($\Delta S_{RT}$) and a term of all other contributions ($\Delta S_{other}$) (Equation 5) (Spolar & Record, 1994). As the entropy change of the association reaction is temperature dependent, a temperature exists where $\Delta S_{assoc}$ equals zero. $T_S$ can be determined by linear regression of $T\Delta S$ as a function of temperature (Figure 4.5).

**Equation 5:** $\Delta S_{assoc}(T_S) = \Delta S_{HE}(T_S) + \Delta S_{RT} + \Delta S_{PE} + \Delta S_{other}$

Because changes in the exposure of nonpolar surface to water are thought to be the major contributor to the heat capacity change of processes involving proteins (Sturtevant, 1977), Baldwin (Baldwin, 1986) proposed a relationship...
in which the contribution of the heat capacity change to the observed entropy change in association reactions could be obtained from the hydrophobic effect ($\Delta S_{\text{HE}}$) (Equation 6) (Spolar & Record, 1994).

**Equation 6:** $\Delta S_{\text{HE}} (T_s) = 1.35 \Delta C_p \ln (T_s/386)$

Since the dependence of $\Delta S_{\text{RT}}$ on temperature and on molecular mass are predicted to be logarithmic, $\Delta S_{\text{RT}}$ should be relatively insensitive to $T_s$ and to the molecular mass of the protein and ligand involved and thus similar for all associations. The magnitude and sign of this term and its insensitivity to the size of the associating species agree well with the statistical mechanical estimate of $\Delta S_{\text{RT}} \approx -50 \text{cal mol}^{-1} \text{K}^{-1}$ for protein-DNA association (Finkelstein & Janin, 1989; Spolar, 1994).

Typical values for $\Delta S_{\text{PE}}$ are between 10 and 60 cal mol$^{-1}$ K$^{-1}$ (Spolar & Record, 1994; Record et al., 1977). The dependence of the equilibrium constants on the salt concentration was measured by EMSA and gave a value of 50 cal mol$^{-1}$ K$^{-1}$ for $\Delta S_{\text{PE}}$ (Künne, 1998; Meierhans, Thesis 1998).

Spolar & Record calculated values of $\Delta S_{\text{HE}}$ at $T_s$ from experimental values of $\Delta C_p$ for folding of globular proteins. They found that the unfavourable entropic contribution expressed per residue was relatively constant for the entire set of proteins they examined [$\Delta S_{\text{other}} = -5.6 \pm 0.5 \text{cal mol}^{-1} \text{K}^{-1}$] (Spolar & Record, 1994). For associative processes, they proposed that $\Delta S_{\text{other}}$ corresponds to the change in conformational entropy on binding (Spolar & Record, 1994). It is assumed to be independent of temperature at least for the small temperature range examined. If conformational changes that involve
folding dominate $\Delta S_{\text{other}}$, then division of $\Delta S_{\text{other}}$ by $-5.6$ cal mol$^{-1}$ K$^{-1}$ yields a thermodynamic estimate of the number of residues $R$ involved in the folding transition (Equation 7).

**Equation 7:** $R = \Delta S_{\text{other}} / -5.6$ cal mol$^{-1}$ K$^{-1}$

### 4.5 Thermodynamic parameters for the binding process

The binding processes between MASH-BHLH and the MCK-S and SP-1 oligonucleotides were strongly exothermic (Tables 3 and 4). In all buffer systems studied, the enthalpies were more negative for binding to E-box-containing DNA than for the association with heterologous DNA sequences. The values measured for $\Delta H$ in low-salt MES and PIPES buffer were indistinguishable within the accuracy of the measurement, indicating that no important protonation or deprotonation step was involved in the binding process (Table 3). Such a change in the protonation state would have led to significantly altered values for $\Delta H$ in the two buffers (Plum & Breslauer, 1995; Ortiz et al., 1998). Similar results were obtained for the DNA binding processess of the TATA binding protein (Petri et al., 1995) and of the transcription factor GCN4 (Berger et al., 1996).

The DNA binding process of MASH-BHLH showed a strong dependence on the reaction temperature of the measured enthalpy $\Delta H$ (Figure 4.5B) and of the derived $T\Delta S$ (Figure 4.5C), which compensate to make $\Delta G$ (Figure 4.5A) almost temperature insensitive, at least within the range studied (Table 4).
Figure 4.5: Thermodynamic profiles of the DNA binding process of MASH-BHLH in low-salt MES buffer as determined by titration calorimetry. 

(A) Plot of the free energy of binding, ΔG, against the temperature for titrations of MCK-S (□) and SP-1 oligonucleotide (▲) with MASH-BHLH. 

(B and C) Plots of the enthalpy ΔH (B) and TΔS (C) against the temperature for titrations of MCK-S (□) and SP-1 oligonucleotide (▲) with MASH-BHLH.
Such behaviour has been consistently observed for specific protein-DNA interactions (Merabet & Ackers, 1995; Ha et al., 1989; Takeda et al., 1992; Murphy & Freire, 1992; Spolar et al., 1992; Spolar & Record, 1994). It is expected for any process which involves water as a solvent and which is characterised by a relatively large ΔC_p (Merabet & Ackers, 1995).

A notable feature of the binding reaction of MASH-BHLH is the fact that T_H, the temperature at which the reaction enthalpy equals zero, is -13°C for binding to specific and non-specific DNA in low-salt MES buffer (Table 4). The value of T_S, the temperature at which the reaction entropy changes sign, is -2°C for E-box-containing DNA and +2°C for heterologous DNA (Figure 4.5C). Many other analyses of specific interactions between proteins and DNA indicate that the change of the driving force of the reaction takes place within the physiological temperature range so that the binding reaction is entropically driven at low temperatures, but enthalpically driven at higher temperatures (Petri et al., 1995; Ha et al., 1989; Foguel & Silva, 1994). The results obtained indicate a different behaviour for the interactions of MASH-BHLH with DNA. Although the entropy changes are favourable at very low temperatures and strongly unfavourable at higher temperatures (Figure 4.5C), the DNA binding reaction of MASH-BHLH is enthalpically driven throughout the physiologically relevant temperature range under the conditions used in this study (Figure 4.5B). A few examples of enthalpically driven associations have been described in the literature (Berger et al., 1996; Ladbury et al., 1994; Merabet & Ackers, 1995).

As mentioned above, the enthalpies of the binding reaction showed a strong dependence on the temperature. In good agreement with the data reported in Figure 4.5B, heat capacity changes were calculated under the assumption that
$\Delta C_p$ was temperature independent in the small temperature range studied. Linear regression of the data in Figure 4.5B gave $\Delta C_p = -733 (\pm 99)$ cal mol$^{-1}$ K$^{-1}$ for the MCK-S oligonucleotide and $\Delta C_p = -575 (\pm 105)$ cal mol$^{-1}$ K$^{-1}$ for the SP-1 oligonucleotide. These values are comparable to the values reported for complex formation between DNA and proteins (Ha et al., 1989; Takeda et al., 1992; Spolar & Record, 1994, Jin et al., 1993; Lundbäck et al., 1993). Large negative $\Delta C_p$ values were first encountered in protein folding. They are now considered typical of specific protein-DNA recognition processes due to the stereospecific interaction of the complementary protein-DNA interface (Ha et al., 1989; Burley, 1994).

4.6 Coupling of MASH-BHLH folding to DNA binding

Changes in the exposure of nonpolar surface area are believed to be the major contributors to the heat capacity changes of DNA binding reactions (Baldwin, 1986; Sturtevant, 1977). However, a more detailed analysis reveals that $\Delta C_p$ also contains contributions from the changes in the water accessibility of the polar surface area (Spolar et al., 1992). Spolar & Record have proposed an equation $[\Delta S_{HE}=0.32 \Delta A_{np} \ln(T_s/386)]$, which is based on the change of nonpolar surface area ($\Delta A_{np}$). This equation is applicable to all protein processes relevant to the entropic contribution of the hydrophobic effect ($\Delta S_{HE}$) from structural information. Where $\Delta A_{np}$ is not known, it is assumed that the ratio $\Delta A_p/\Delta A_{np}$ is 0.59, as in the folding of globular proteins - an assumption which is justified in many binding reactions (Spolar et al., 1992; Spolar & Record, 1994; Berger et al., 1996). An empirical correlation between
\( \Delta C_p \) and the entropy change of the binding reaction due to the hydrophobic effect, \( \Delta S_{HE} \), could be derived (Spolar et al., 1992) (Eq. 6). \( \Delta S_{HE} \) for the reaction between MASH-BHLH and the MCK-S oligonucleotide was calculated to be 357 (±48) cal K\(^{-1}\) mol\(^{-1}\). Due to the uncertainty in the ratio of the polar and the nonpolar water-accessible surface area, the error introduced in \( \Delta S_{HE} \) is likely to be approximately 10\% (Spolar & Record, 1994). However, it must be kept in mind that the direct calculation of \( \Delta S_{HE} \) is also rather inaccurate mainly due to the significant error introduced by determination of the surface area from structural data.

The unfavourable entropic term \( \Delta S_{RT} \) resulting from the reduction in the available rotational and translational degrees of freedom of MASH-BHLH and the oligonucleotide upon association was assumed to be independent of \( T_s \) and the molecular mass of the reactants. A value of -50 cal K\(^{-1}\) mol\(^{-1}\) was used (Spolar & Record, 1994; Finkelstein & Janin, 1989). The conformation of the DNA is not changed significantly upon binding to BHLH proteins such as MASH-1, and it remains essentially in the B-form (Ma et al., 1994; Ellenberger et al., 1994; Ferré-D’Amaré et al., 1993; Ferré-D’Amaré et al., 1994; Meierhans et al., 1995; Künne et al., 1996; Allemann & Egli, 1997). Based on these assumptions, the entropy change resulting from local protein folding transitions coupled to DNA binding and the number of residues involved in the folding transition could be calculated to be 307 (±48) cal mol\(^{-1}\) K\(^{-1}\) and 54, respectively (Spolar & Record, 1994). MASH-BHLH binds to DNA as a symmetrical dimer (Meierhans et al., 1995), and therefore 27 amino acid residues of each subunit undergo a transition from an unordered to a well-defined conformation upon DNA binding. Inspection of Figure 2.3 reveals that the N-terminal 27 amino acids comprise the entire basic domain, which is
known to interact directly with DNA (Ma et al., 1994; Ellenberger et al., 1994), and the two N-terminal amino acids of helix-1. Because we have neglected the contribution of the polyelectrolyte effect, $\Delta S_{PE}$, to the observed entropy change (Record et al., 1976; Record et al., 1991), the value reported might represent an upper limit for the actual number of residues undergoing a folding transition. Typical values for $\Delta S_{PE}$ are between 10 and 60 cal mol$^{-1}$ K$^{-1}$ (Spolar & Record, 1994). Measurements of the dependence of the equilibrium constant on the salt concentration by EMSA gave a value of 50 cal mol$^{-1}$ K$^{-1}$ for $\Delta S_{PE}$ in low salt buffer (Record et al., 1977), resulting in an decrease of approximately 15% in the number of amino acid residues which change conformation upon complex formation.

In agreement with these data, nuclear magnetic resonance studies of the BHLH protein E47 showed that the basic region was in a random coil conformation in the absence of DNA even at the millimolar concentrations used in the NMR experiments (Fairman et al., 1997). The structure of the helix-loop-helix domain of the E47 homodimer, on the other hand, was preserved in solution when compared to the co-crystal structures of E47 and DNA (Ellenberger et al., 1994).

### 4.7 CD spectroscopy studies of the binding process

The nature of the change in the conformation of MASH-BHLH upon DNA binding was also studied by circular dichroism spectroscopy. The CD spectrum of a 100 $\mu$M solution of MASH-BHLH showed that in the absence of DNA, the protein adopted a conformation which was approximately 50%
alpha-helical and 50% random coil (Figure 4.6) (Greenfield & Fasman, 1969; Compton & Johnson, 1986). Increasing the MASH-BHLH concentration did not change the CD spectrum. According to X-ray analyses of the DNA complexes of MyoD and E47, the unstructured loop connecting helix 1 and helix 2 consists of approximately eight residues (Ma et al., 1994; Ellenberger et al., 1994). Therefore, an additional 16 of the 68 amino acids of MASH-BHLH were in an unstructured conformation in the absence of DNA.

The CD spectrum of equimolar amounts of the MASH-BHLH dimer and double-stranded MCK-S oligonucleotide revealed a conformational change in part of the protein, in that the amount of alpha-helicity was increased to approximately 85%, or 58 of 68 amino acid residues (Figure 4.6). Similar results were obtained for the addition of 1 equivalent of SP-1 oligonucleotide to a solution of MASH-BHLH. This observation could be explained by the formation of an alpha-helix over about 25 residues in the basic region of MASH-BHLH upon DNA binding. It agrees well with the results obtained from the ITC experiments.
Figure 4.6: Characterisation of the conformation of MASH-BHLH by CD spectroscopy. Curve a, 50 μM MCK-S oligonucleotide (no protein present); curve b, 100 μM MASH-BHLH; curve c, mixture of 50 μM MCK-S and 100 μM MASH-BHLH (equivalent to 50 μM MASH-BHLH dimer). The contribution from the free oligonucleotide (curve a) was subtracted from the spectrum of the complex.

Earlier results had shown that MASH-BHLH was largely unfolded at the concentrations where DNA binding occurs but adopted a mainly α-helical structure in the presence of DNA (Meierhans et al., 1995; Künne & Allemann, 1997). Similarly, at least part of the BHLH domain was found to fold upon dimerisation (Künne & Allemann, 1997), and under the conditions of the ITC experiment, MASH-BHLH was found to be mainly dimeric with dimerisation mediated through the stably folded helix-loop-helix domain. However, the experiments described above indicate that the basic region is unfolded even at the high concentrations used in ITC experiments. In this context, it is worthy
of note that unfavourable interactions are formed between the positive charges of the side chains of arginines 119, 120, 123, and 127 (Figure 4.7) when the basic region of MASH-BHLH forms an alpha-helix in the absence of DNA.

**Figure 4.7:** Forming an alpha-helical structure without DNA would cause electrostatically unfavourable interactions between the positively charged residues in the basic regions.

However, according to the X-ray structures of the DNA complexes of MyoD and E47, the $\alpha$-helicity of the basic region allows these arginines residues to form stabilising interactions with the DNA phosphate backbone, thereby mitigating the repulsive interactions between them.
Chapter 5: Thermodynamics of MM17/DNA complex formation

5.1 Introduction

CD-spectroscopic experiments with the BHLH proteins MASH-1 and MyoD have shown that in the absence of DNA, BHLH proteins can form stable dimers, which are in a concentration-dependent equilibrium with the monomer (Meierhans et al., 1995; Künne et al., 1996). Dimerisation is accompanied by a folding transition from a largely unfolded monomer to a predominantly alpha-helical dimer. Isothermal titration calorimetry (ITC) and CD-spectroscopy have indicated that the basic region remains unfolded in the dimer, but adopts an alpha-helical conformation on DNA binding (Künne et al., 1998). Folding and dimerisation of the BHLH domains can be induced through the addition of DNA, even at concentrations where the BHLH domain alone is mainly unfolded.

The construction of MM17, a single chain dimer of MASH-BHLH, in which the C-terminus of the BHLH subunit is attached to the N-terminus of the second subunit through a peptide linker of 17 amino acids is reported in chapter 3. Characterisation by electrophoretic mobility shift assay and CD-spectroscopy revealed that MM17 and wild type MASH-BHLH displayed similar DNA recognition properties. However, in MM17 the hydrophobic core between alpha-helices-1 and -2 of the two HLH domains was formed even in

---

the absence of DNA. "Dimerisation" was independent of the concentration because of the monomeric nature of MM17. To complement the structural and biochemical studies, the thermodynamic parameters of the DNA binding reaction of MM17 were measured by ITC.

5.2 ITC Measurements

The thermodynamic parameters for the association process between MM17 and the MCK-S and SP-1 oligonucleotides were determined by isothermal titration calorimetry experiments in the range from 12°C to 32°C (Table 5 and Figure 5.1). For all titrations, the best fit of the experimental data to the binding isotherm was obtained for one MM17 molecule binding to one double stranded DNA oligonucleotide. This is expected because the two BHLH subunits are covalently linked in MM17 (Figure 5.1).

In low salt buffer at pH 7 the association constants for MM17 binding to MCK-S and SP-1 were 1.11 (±0.11) x 10⁷ M⁻¹ and 5.04 (±2.75) x 10⁶ M⁻¹ at 12°C and 4.38 (±1.87) x 10⁷ M⁻¹ and 6.58 (±2.97) x 10⁶ M⁻¹ at 32°C, respectively. The free energy required to transfer MM17 from an E-box to a heterologous DNA sequence was 0.5 kcal/mol at 12°C (Table 5) which is in good agreement with the value determined in EMSA experiments in chapter 3 (Table 2). The specificity of DNA binding of MM17 increased with increasing temperature (Figure 5.2) and reached a ΔΔG of 1.15 kcal/mol at 32°C (Table 5). These specificities were similar to those measured by ITC for wild-type MASH-BHLH (Chapter 4).
Figure 5.1: Example of an ITC experiment. MCK-S DNA was titrated with increasing amounts of MM17 in low-salt PIPES buffer at pH 7 and 27°C.
Table 5: Thermodynamic Parameters of the Binding Reaction of MM17 and Different Binding Sites at pH 7.0

<table>
<thead>
<tr>
<th>T (K)</th>
<th>$K_A$ (M$^{-1}$)</th>
<th>$\Delta G$ (kcal/mol)</th>
<th>$\Delta H$ (kcal/mol)</th>
<th>$T \Delta S$ (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MCK-S in low-salt PIPES Buffer pH 7.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>285</td>
<td>1.11(±0.11)E+07</td>
<td>-9.18(±0.05)</td>
<td>-35.71(±1.87)</td>
<td>-26.53(±1.93)</td>
</tr>
<tr>
<td>290</td>
<td>1.77(±0.65)E+07</td>
<td>-9.58(±0.24)</td>
<td>-42.78(±3.94)</td>
<td>-33.19(±4.17)</td>
</tr>
<tr>
<td>295</td>
<td>6.04(±8.65)E+07</td>
<td>-10.41(±0.73)</td>
<td>-36.89(±2.95)</td>
<td>-26.48(±3.31)</td>
</tr>
<tr>
<td>300</td>
<td>2.40(±1.16)E+07</td>
<td>-10.07(±0.35)</td>
<td>-44.06(±0.83)</td>
<td>-33.99(±1.04)</td>
</tr>
<tr>
<td>305</td>
<td>4.38(±1.87)E+07</td>
<td>-10.61(±0.29)</td>
<td>-49.42(±1.91)</td>
<td>-38.81(±1.87)</td>
</tr>
<tr>
<td></td>
<td>SP-1 in low-salt PIPES Buffer pH 7.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>285</td>
<td>5.04(±2.75)E+06</td>
<td>-8.68(±0.32)</td>
<td>-28.63(±1.13)</td>
<td>-19.95(±0.88)</td>
</tr>
<tr>
<td>290</td>
<td>8.98(±3.17)E+06</td>
<td>-9.19(±0.22)</td>
<td>-32.53(±2.72)</td>
<td>-23.33(±2.90)</td>
</tr>
<tr>
<td>295</td>
<td>1.70(±0.86)E+07</td>
<td>-9.69(±0.33)</td>
<td>-31.39(±3.87)</td>
<td>-21.69(±4.20)</td>
</tr>
<tr>
<td>300</td>
<td>1.33(±0.73)E+07</td>
<td>-9.84(±0.32)</td>
<td>-37.75(±5.88)</td>
<td>-27.91(±6.20)</td>
</tr>
<tr>
<td>305</td>
<td>6.58(±2.97)E+06</td>
<td>-9.46(±0.34)</td>
<td>-39.79(±6.38)</td>
<td>-30.33(±6.05)</td>
</tr>
<tr>
<td></td>
<td>MCK-S in 25 mM PIPES Buffer pH 7.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>285</td>
<td>4.07(±1.66)E+06</td>
<td>-8.58(±0.26)</td>
<td>-32.56(±3.64)</td>
<td>-23.98(±3.65)</td>
</tr>
<tr>
<td>290</td>
<td>2.01(±2.13)E+07</td>
<td>-9.48(±0.59)</td>
<td>-39.71(±28.6)</td>
<td>-30.23(±28.0)</td>
</tr>
<tr>
<td>295</td>
<td>3.99(±1.60)E+06</td>
<td>-8.88(±0.22)</td>
<td>-37.64(±7.91)</td>
<td>-28.76(±8.07)</td>
</tr>
<tr>
<td>300</td>
<td>2.57(±1.17)E+06</td>
<td>-8.76(±0.28)</td>
<td>-34.88(±17.9)</td>
<td>-26.12(±18.2)</td>
</tr>
<tr>
<td>305</td>
<td>2.58(±2.02)E+06</td>
<td>-8.83(±0.53)</td>
<td>-29.18(±6.63)</td>
<td>-20.34(±6.09)</td>
</tr>
<tr>
<td></td>
<td>SP-1 in 25 mM PIPES Buffer pH 7.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>285</td>
<td>3.20(±1.31)E+06</td>
<td>-8.45(±0.22)</td>
<td>-32.24(±6.42)</td>
<td>-23.79(±6.64)</td>
</tr>
<tr>
<td>290</td>
<td>1.71(±0.93)E+06</td>
<td>-8.22(±0.33)</td>
<td>-43.70(±2.25)</td>
<td>-35.48(±2.58)</td>
</tr>
<tr>
<td>300</td>
<td>2.49(±0.00)E+06</td>
<td>-8.78(±0.00)</td>
<td>-26.66(±0.00)</td>
<td>-17.88(±0.00)</td>
</tr>
</tbody>
</table>
Figure 5.2: Thermodynamic profiles of the DNA binding processes of MM17 in low-salt PIPES buffer at pH 7 as determined by ITC. Plots of ΔG, ΔH and TΔS as a function of temperature for titrations of MM17 with MCK-S (○) and SP-1 (■). The heat capacity change, ΔC_p, was calculated by linear regression from the values of ΔH = ΔH(T).
5.3 pH dependence of the binding process

The effect of pH on the stability of the DNA complexes of MM17 was measured by ITC for the pH range between 6 and 8 (Table 6). These measurements showed that within experimental error the complex stability and therefore the DNA binding specificity were independent of pH at least for the pH range examined (Figure 5.3). No important protonation or deprotonation step occurred during the association reactions of MM17 and DNA. This finding was not surprising, since the protonation state did not change in the binding reaction of wild-type MASH-BHLH (Chapter 4).

Table 6: Dependence of the Thermodynamic Parameters of the Binding Reaction of MM17 on Buffer and pH at 290 K.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH</th>
<th>ΔG (kcal/mol)</th>
<th>ΔH (kcal/mol)</th>
<th>TΔS (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCK-S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MES</td>
<td>6.0</td>
<td>-8.9(±0.15)</td>
<td>-41.9(±3.28)</td>
<td>-32.9(±3.26)</td>
</tr>
<tr>
<td>PIPES</td>
<td>7.0</td>
<td>-9.6(±0.24)</td>
<td>-42.8(±3.94)</td>
<td>-33.2(±4.17)</td>
</tr>
<tr>
<td>TRIS</td>
<td>8.0</td>
<td>-9.4(±0.33)</td>
<td>-23.1(±5.41)</td>
<td>-13.7(±0.57)</td>
</tr>
<tr>
<td>SP-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MES</td>
<td>6.0</td>
<td>-8.9(±0.08)</td>
<td>-39.2(±3.42)</td>
<td>-30.2(±3.50)</td>
</tr>
<tr>
<td>PIPES</td>
<td>7.0</td>
<td>-9.2(±0.22)</td>
<td>-32.5(±2.72)</td>
<td>-23.3(±2.90)</td>
</tr>
<tr>
<td>TRIS</td>
<td>8.0</td>
<td>-8.8(±0.19)</td>
<td>-25.1(±6.45)</td>
<td>-16.3(±6.64)</td>
</tr>
</tbody>
</table>
5.4 Thermodynamic parameters for the binding process

The binding processes between MM17 and MCK-S and SP-1 were strongly exothermic in the temperature range examined (Table 5). The reaction was more exothermic for binding to E-box-containing DNA. The reaction enthalpy ΔH and the reaction entropy TΔS, both of which are functions of the reaction temperature, compensate to make the reaction free energy almost independent of temperature. This behaviour mirrors that observed for wild-type MASH-BHLH binding to DNA and is a general feature of bimolecular association reactions in aqueous solvents (Merabet & Ackers, 1995). $T_s$, the temperature at which the entropy change of the association reaction of MM17 and MCK-S
changes sign, was 232 K. The enthalpy change for MCK-S binding of MM17 was zero for the temperature $T_\text{m} = 222$ K. As a consequence, the association reaction is enthalpy driven over the whole physiological temperature range. Due to the strong temperature dependence of the reaction enthalpy, a large negative change of the heat capacity was measured. Linear regression of the data in Figure 5.2 gave a $\Delta C_p$ of $-574 \pm 239$ cal mol$^{-1}$ K$^{-1}$ for the E-box-containing oligonucleotide and $-550 \pm 113$ cal mol$^{-1}$ K$^{-1}$ for the heterologous SP-1 oligonucleotide. These values are comparable to those measured for the DNA binding processes of MASH-BHLH and contain not just the contributions from the docking of pre-ordered structures but also from any conformational changes that occur on DNA binding of MM17.

5.5 Interpretation of the observed heat capacity change

The large negative heat capacity changes typically observed for the formation of specific protein-DNA complexes are caused to a large extent by changes in the exposure of nonpolar surface area (Baldwin, 1986; Takeda et al., 1992; Ha et al., 1989; Burley, 1994; Spolar & Record, 1994; Lundbäck et al., 1993; Jin et al., 1993; Berger et al., 1996; Künne et al., 1998). A more detailed analysis revealed however, that changes in the exposure of nonpolar surface area also contribute to $\Delta C_p$ (Spolar et al., 1992; Spolar & Record, 1994). Calculation of $\Delta S_{\text{HE}}$ from the measured values of $\Delta C_p$ gave $\Delta S_{\text{HE}}$ of $394(\pm 164)$ cal mol$^{-1}$ K$^{-1}$ for MCK-S and $335(\pm 69)$ cal mol$^{-1}$ K$^{-1}$ for SP-1 binding, respectively (see
Because one protein molecule reacts with one double stranded oligonucleotide to form the complex a value of -50 cal mol$^{-1}$ K$^{-1}$ was used for $\Delta S_{RT}$ (Künne et al., 1998; Finkelstein & Janin, 1989).

Due to the low solubility of MM17 in buffers containing even modest salt concentrations no reliable value for the contribution of the temperature independent polyelectrolyte effect, $\Delta S_{PE}$, to the observed entropy change could be measured directly (Record et al., 1976; Record et al., 1991; Anderson & Record, 1995). However, CD-spectroscopy has shown that the DNA complexes of MM17 and MASH-BHLH are very similar, so a similar $\Delta S_{PE}$ was assumed. $\Delta S_{PE}$ had been estimated as 50 cal mol$^{-1}$ K$^{-1}$ for the DNA binding reaction of MASH-BHLH in EMSA experiments (Record et al., 1976; Record et al., 1991; Künne et al., 1998; Meierhans & Allemann, 1998).

CD-spectroscopy had indicated that the conformation of the DNA did not change significantly during the DNA binding processes of MM17 and MASH-BHLH (Sieber & Allemann, 1998; Meierhans et al., 1995; Meierhans & Allemann, 1998; Künne et al., 1998). All available X-ray structures of DNA complexes of BHLH proteins indicate that the DNA is essentially in its B-form in the complex (Ma et al., 1994; Ellenberger et al., 1994). Therefore, no contributions to the overall entropy change of the association reaction is expected from a conformational change of the DNA. The remaining entropy change $\Delta S_{conf} = -(\Delta S_{HE} + \Delta S_{PE} + \Delta S_{RT})$, which was calculated as -394 cal mol$^{-1}$ K$^{-1}$, was attributed to local protein folding transitions coupled to DNA binding. Under the experimentally (Spolar & Record, 1994) and computationally (Dill, 1985; Alonso & Hutchinson, 1989; Creamer & Rose,
1992; Picket & Sternberg, 1993) well-justified assumption that for every residue which undergoes a folding transition 5.6 cal mol\(^{-1}\) K\(^{-1}\) on average are lost, 70 residues, which are unordered in free MM17, adopt a folded form on DNA binding. CD-spectroscopy had revealed that helix-1 and helix-2 of both HLH domains in MM17 adopted an alpha-helical structure even in the absence of DNA (Figure 3.14). However, the basic regions of MM17 underwent a conformational change on DNA binding from a largely unstructured to a mainly alpha-helical structure. Similarly, the analysis of the thermodynamic data obtained for the DNA binding reaction of MASH-BHLH revealed that 27 amino acids per subunit underwent a conformational change on DNA binding, indicating that the whole basic region and the N-terminal two amino acids of helix-1 change conformation (see Chapter 4). The ITC data obtained for the associations of MM17 with DNA indicate that in addition to the 54 amino acids of the two basic regions another 16 amino acids lose conformational degrees of freedom in the complex. According to the crystal structure of the DNA complexes of MyoD (Ma et al., 1994) and E47 (Ellenberger et al., 1994), the distance between the N-terminal end of one subunit and the C-terminal end of the other is approximately 50 Å. This distance can only just be spanned by a linker of length 18 (Figure 5.4) and several of the residues at the N-terminal end of the basic region will not be in alpha-helical conformation but will be part of the linker region. The analysis of the thermodynamic data indicates that due to its short length the linker region of MM17 shows little flexibility and adopts a well-defined structure in the DNA-complex.
Figure 5.4: Cartoon of the connectivities between individual BHLH subunits of MASH-1 to produce the MM17 protein. The amino acid sequences of the BHLH domain and of the linker are shown.
Chapter 6. Conclusions

Designing a different network of contacts with novel DNA binding specificity revealed that two strictly conserved amino acids are central for the DNA binding specificity displayed by BHLH proteins. An arginine which does not contact the nucleobases locks the conformation of a glutamate. In its locked conformation, the carboxylate of the glutamate interacts favourably only with E-box containing DNA, while repulsive interactions destabilise complexes with all other DNA sequences. Although affinity was retained, all mutations in these positions led to proteins with reduced binding specificity.

A thermodynamic analysis of BHLH/DNA complex formation has given an insight into the binding process in terms of energies. Interpretation of the heat capacity change and dissection of the entropy change has shown the protein undergoes a folding transition upon DNA binding. This agreed well with CD measurements of wild-type MASH-BHLH and the single chain 'dimers'. We conclude that the subunits have dimerised at the high concentrations used in the experiments, but the basic regions are unfolded and adopt an alpha-helical conformation only upon DNA binding.

Covalently linking the two subunits of the BHLH dimer increased the affinity of interaction with DNA and modestly increased specificity. The single chain proteins thus preserve most of the characteristic DNA binding properties of wild-type MASH-BHLH. They also provide an opportunity to address several questions concerning molecular recognition. Since amino acids in the two domains can be varied independently, it should be possible through mutagenesis to direct the single chain proteins to assymmetric DNA target
sequences. In addition, single chain proteins can be displayed on the surface of filamentous phage particles so that new DNA binding properties can be selected through random mutagenesis from a large repertoire of mutant proteins.
# Chapter 7: Materials and Methods

## 7.1 Materials

### Chemicals

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Abbreviation</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-Bromo-4-chloro-3-indolyl-phosphate</td>
<td>BCIP</td>
<td>Sigma</td>
</tr>
<tr>
<td>Acetic acid</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Acrylamide</td>
<td></td>
<td>Sigma</td>
</tr>
<tr>
<td>Agarose electrophoresis grade</td>
<td></td>
<td>GIBCO BRL</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>NH$_4$OAc</td>
<td>Fluka</td>
</tr>
<tr>
<td>Ammonium peroxodisulfate</td>
<td>AP</td>
<td>Sigma</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>(NH$_4$)$_2$SO$_4$</td>
<td>Fluka</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>amp</td>
<td>Sigma</td>
</tr>
<tr>
<td>Bacto Tryptone</td>
<td></td>
<td>DIFCO</td>
</tr>
<tr>
<td>Boric acid</td>
<td>H$_3$BO$_3$</td>
<td>Fluka</td>
</tr>
<tr>
<td>Bromophenol Blue</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Butanol</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Calcium chloride</td>
<td>CaCl$_2$</td>
<td>Fluka</td>
</tr>
<tr>
<td>3-[(3-Cholamidopropyl)-dimethylammonio]-1-propanesulfate</td>
<td>CHAPS</td>
<td>Sigma</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>cl-amp</td>
<td>Fluka</td>
</tr>
<tr>
<td>Coomassie Brilliant Blue</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>DTT</td>
<td>Fluka</td>
</tr>
<tr>
<td>Ethylenediaminetetraacetic acid</td>
<td>EDTA</td>
<td>Fluka</td>
</tr>
<tr>
<td>Ethanol</td>
<td>EtOH</td>
<td>Fluka</td>
</tr>
<tr>
<td>Ethidium bromide</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Formamide</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Glucose</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Glycerol</td>
<td></td>
<td>Fluka</td>
</tr>
<tr>
<td>Isopropyl-β-D-thiogalactopyranoside</td>
<td>IPTG</td>
<td>Sigma</td>
</tr>
<tr>
<td>Magnesium chloride</td>
<td>MgCl$_2$</td>
<td>Fluka</td>
</tr>
<tr>
<td>Manganese chloride</td>
<td>MnCl$_2$</td>
<td>Fluka</td>
</tr>
<tr>
<td>2-Mercaptoethanol</td>
<td></td>
<td>Sigma</td>
</tr>
<tr>
<td>2-Morpholinoethanesulfonic acid</td>
<td>MES</td>
<td>Fluka</td>
</tr>
</tbody>
</table>
Methanol Fluka
3-Morpholinopropanesulfonic acid MOPS Fluka
N,N'-methylene-bis-acrylamide Fluka
Nitroblue tetrazolium NBT Sigma
Piperazine-1,4-bis(2-ethanesulfonic acid) PIPES Fluka
Phenol/Chloroform/Isoamylalc. (49.5:49.5:1) Fluka
Phenylmethylsulfonyl fluoride PMSF Fluka
Phosphoric acid H₃PO₄ Fluka
Polyethylene glycol PEG 6000 Fluka
Potassium acetate KOAc Fluka
Potassium hydroxide KOH Siegfried Handel
Red Ponceau Sigma
Ribonuclease A RNase A Sigma
Rubidium chloride RbCl Fluka
Sodium acetate NaOAc Fluka
Sodium chloride NaCl Fluka
Sodium hydroxide NaOH Siegfried Handel
Sodium dodecyl sulfate SDS Fluka
N,N,N',N'-Tetramethylethylenediamine TEMED Fluka
Tris(hydroxymethyl)-aminomethane Tris Sigma
Urea Riedel de Haen, Anachem
Xylene cyanol Fluka
Yeast Extract DIFCO

Media (Sambrook et al., 1989)

LB: 10 g Bacto Tryptone, 5 g Yeast Extract, 10 g Sodium chloride
2 x YT: 16 g Bacto Tryptone, 10 g Yeast Extract, 10 g Sodium chloride

Amounts are for 1 litre of liquid media. All media were autoclaved immediately after preparation and stored at 4°C. For agar plates, 15 g of Bacto Agar were added to LB medium. After autoclaving, the solution was left to
cool to about 40°C, at which point antibiotics were added. The mixture was swirled gently and poured into petri dishes (Ø 9 cm). The plates were stored bottom up in a plastic bag at 4°C.

**Buffers**

Buffers were made as stock solutions according to Maniatis (Sambrook et al., 1989).

50 x TE: 10 mM Tris-Cl (pH 8.0), 1mM EDTA (pH 8.0)
5 x TBE: 54 g Tris base, 27.5 g boric acid, 20 ml 0.5 M EDTA (pH 8.0)
50 x TAE: 242 g Tris base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA (pH 8.0)

**E. coli strains**

DH5α  
*endA1 hsdR17 (rK- mK+) supE44 thi-1 recA1 gyrA(NaI') relA1 Δ(lacIZYA-argF)U169 deoR (ϕ80dlacΔ(lacZ)M15); F-

BL21(DE3)pLysS  
dcm ompT hsdS(rK- mK+) gal λ(DE3) [pLysS Cam']; F

MC1061  
hsdR mcrB araD139 Δ(araABC-leu)7679 ΔlacI74 
galU galK strA thi; F

XL1 blue  
*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacF'ZΔM15 Tn10 (Tet')]

CJ236  
dut, ung, thi, relA; pCJ105 (Cm')
7.2 Methods

Preparation of Competent Cells

An *E.coli* strain was streaked out on an agar plate containing no antibiotics and then left to grow overnight at 37°C. Two colonies were picked and each grown overnight in 5 ml of LB-Medium. Each culture was then grown in 250 ml of LB-Medium (37°C, shaken at 300 rpm) until an OD$_{600}$ value of 0.5 had been reached. A Varian DMS 300 UV/VIS Spectrophotometer was used to measure the absorbance. The cells were cooled on ice and centrifuged in a pre-cooled rotor (GS3) of a Sorvall centrifuge for 15 min at 5000 rpm and 4°C. The rest of the procedure was carried out in the cold room (4°C). The cells were suspended in ice-cold Buffer 1 (100 mM RbCl, 50 mM MnCl$_2$, 30 mM KOAc, 10 mM CaCl$_2$, 15% (v/v) glycerol) for competent cells (30 ml buffer/100 ml culture). The newly suspended cells were centrifuged (10 min, 5000 rpm, GS3 rotor) and then resuspended in Buffer 2 (10 mM MOPS pH 7, 10 mM RbCl, 75 mM CaCl$_2$, 15% (v/v) glycerol) for competent cells (4 ml buffer/100 ml culture). The cells were finally pipetted into Eppendorf tubes (300 µl aliquots) which were frozen in liquid N$_2$ and then stored at -80°C.

Transformation

Competent cells were thawed on ice. 100 µl of competent cells and 1 µl of the plasmid were added to an eppendorf tube and left for 20 min on ice, 2 min at 42°C, and 2 min on ice again. LB-Medium (400 µl) was added to this
suspension and incubated for 1 hour at 150 rpm and 37°C. The solution was centrifuged for 20 seconds at 14'000 rpm in an Eppendorf centrifuge and the supernatant was discarded. Fresh LB-medium (100 µl) and 1 µl 50 mg/ml of the appropriate antibiotic was added to the pellet. The cells were resuspended and pipetted onto a pre-dried agar plate containing antibiotic. The plate was incubated overnight at 37°C.

For transformation of ligation products, the procedure was as described above except that 150 µl of competent cells were added to the whole ligation reaction (20 µl).

**DNA Purification and Isolation**

**a) DNA Minipreps**

Cells were grown in 5 ml of LB medium containing the appropriate antibiotic at 37°C and 250 rpm until saturation was reached (normally overnight). The culture was spun down sequentially into an Eppendorf tube and the supernatant discarded. The pellet was resuspended in 100 µl of cold Solution I (50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0) by vortexing. 200 µl of a solution containing 0.2 M sodium hydroxide and 1% SDS was added. The solutions were mixed by gentle inversion of the tube and left for 5 min at room temperature. Chromosomal DNA was precipitated by the addition of 300 µl Solution III (3 M potassium acetate, 2 M glacial acid) and incubated on ice for 5 min. The chromosomal DNA was pelleted by centrifuging in an Eppendorf centrifuge for 7 min at room temperature. The supernatant was transferred to a new tube, incubated with 3 µl RNase A (50 mg/ml) and
extracted twice with 300 μl phenol/chloroform mixture (chloroform: phenol: isoamylalcohol 49.5:49.5:1) and twice with chloroform (chloroform: isoamylalcohol 24:1 v/v). The DNA was precipitated by the addition of 1/10 volume of 3 M sodium acetate (pH 5.0) and 2.5 volumes of ethanol, and incubated at -20°C for 30 min. After centrifuging for 10 min at 4°C, the pellet was washed with 200 μl of ice cold 70% ethanol, dried and resuspended in 20 μl of 1x TE.

b) Wizard® Plus SV Miniprep

The Wizard® Plus SV Miniprep procedure from Promega® was used for DNA purification and isolation. Cells were transformed with a plasmid and grown at 37°C overnight on an agar plate containing ampicillin. LB medium (5 ml) was added to a sterile plastic tube the next afternoon, a medium sized colony picked from the plate and grown in the liquid medium overnight. A pellet was isolated from 5 ml of cell culture by repeatedly centrifuging for 20 seconds at 14000 rpm in an Eppendorf centrifuge and then discarding the supernatant. Wizard® Resuspension solution (250 μl) was added and the pellet vortexed well. Wizard® Lysis solution (250 μl) was added and carefully mixed by inversion of the tube until the solution became clear. Basic Protease solution (8 μl) was added and carefully mixed for 5 minutes at room temperature. Next, Wizard® Neutralisation solution (350 μl) was added and carefully mixed again. The lysate was centrifuged for 10 min at 14000 rpm and room temperature and then loaded onto a Wizard®-Column. The Wizard®-Column was centrifuged for 1 min at 14000 rpm at room temperature and the filtered solution was discarded. Wizard® Wash solution (750 μl) was added to the
column and centrifuged for 1 min at 14000 rpm and room temperature. The filtered solution was discarded. Wizard® Wash solution (250 μl) was added and the column was centrifuged for 2 min at 14000 rpm and room temperature. The Wizard®-Column was transferred to a new Eppendorf tube. The DNA was eluted with 50 μl of Nuclease-free water and centrifuged for 1 min at 14000 rpm in an Eppendorf centrifuge at room temperature. The DNA-solution was stored at -20°C.

**Digestion with restriction enzymes**

Digestion reactions were performed with enzymes either from New England Biolabs or Pharmacia in their recommended buffers. Total volume for the reactions was 20 μl.

**Ligations**

Reactions were carried out in a volume of 20 μl. 2 μl of Ligase Buffer (note that not all supplied buffers contain ATP!) and 1 μl T4 DNA Ligase (New England Biolabs) were added. Reactions were either incubated for 4 hours at 16°C or overnight at room temperature. Digested vector alone was ligated as a control.
Reactions with modifying enzymes

T7 polymerase (USB Amersham)
T4 Polynucleotide Kinase (Pharmacia)
Taq Polymerase (Pharmacia)
Reaction volumes were 20 μl.

Sequencing

All sequencing was done according to the dideoxy nucleosidetriphosphate chain termination method of Sanger (Sanger et al., 1977) directly from the plasmid. The Sequenase Version 2.0 Kit (Amersham - USB) was used.

6 μl plasmid DNA from a Wizard® DNA Minipreparation (Promega) was added to a 1.5 ml microtube and denatured with 4 μl of 1 M NaOH. This solution was incubated at room temperature for 10 min. 3 μl of 3 M NaOAc pH 5.0, 7 μl of water and 75 μl of ethanol were added and the DNA precipitated at -20°C for 30 min. The tube was centrifuged in an Eppendorf centrifuge for 10 min at 4°C, the supernatant was discarded, the pellet washed with 80 μl of 70% ice-cold ethanol, and dried on the bench by placing the tube upside down on a papertowel. The DNA was resuspended in 7 μl of water and 1 μl of primer (16 pmol/μl, ~100 ng) and 2 μl of sequenase buffer (200 mM Tris-HCl pH 7.5, 100 mM MgCl₂, 250 mM NaCl) were added. The tube was transferred to a 65°C water bath and kept at that temperature for 2 min. The solution was then left to cool slowly to room temperature. To the annealed primer-template solution 1 μl of 0.1 M DTT, 2 μl of labelling mix (1.5 μM of
each dNTP), 0.5 µl α-[35S]-dATP (1000 Ci/mmol, 10 µCi/µl), and 1.5 µl of
diluted Sequenase (3 units) were added. After incubation for 5 min at room
temperature, 3.5 µl of this solution were added to each tube containing 2.5 µl
of one of the four dideoxy termination mixes prewarmed to 37°C. They were
then all incubated for 5 min at 37°C.

ddG Termination Mix:  80 µM of each dNTP, 8 µM ddGTP, 50 mM NaCl
The other termination mixes are analogous.

The reaction was stopped by addition of 4 µl stop solution (95% formamide,
20 mM EDTA, 0.05% bromophenol blue, 0.02% xylene cyanol) to each of the
four termination reactions. The samples were heated to 80°C for 2 minutes
followed by a quick transfer to ice immediately before they were loaded onto
the gel.

**Sequencing gels**

The sequencing reactions were analysed on a denaturing polyacrylamide gel
(6% acrylamide/N,N'-bisacrylamide (30:1), 7 M urea). 10.2 g of acrylamide,
0.34 g of bis-acrylamide, and 71.5 g of urea were dissolved in 60 ml of water.
Then, 136 ml water and 2 g of Mixed Bed Resin (BioRad) were added. 34 ml
5x TBE buffer and the acrylamide solution were filtered and degassed, 100 µl
TEMED and 170 µl of a 20% solution of ammonium peroxydisulfate added,
and poured between 2 glass plates (50.4x 33x 0.6 cm back, 48.5x 33x 0.6 cm
front) separated by 0.2 mm spacers. The gels were left overnight to
polymerise. The gels were pre-run for 500 V/H at 80 mA. The samples were loaded, and the gel was run at 300 mA until the gel temperature reached 45 to 50°C. Then, the gel was run at 80 mA for 4500 V/H. 'Smiling' of the gels could be suppressed by attaching an aluminium plate to the glass plates for an even heat distribution. One glass plate was taken away and the urea removed by transferring the gel sitting on the other glass plate into a bath containing a solution of 15% methanol and 10% acetic acid in water. After 45 min the gel was transferred to Whatman 3 MM paper and dried on a Model 583 gel dryer (BioRad) at 80°C. An X-ray film (Kodak BioMax) was exposed to the dried gel overnight at -80°C.

Agarose gels

Agarose gels were used for analysis and isolation of large (<200 bp) pieces of DNA. Agarose was melted in 1xTBE buffer (45 mM Tris-borate, 1 mM EDTA) in a microwave oven to a final concentration of 1-1.8%, depending on the size of the DNA to be analysed. 10 µl of ethidium bromide solution (10 mg/ml) was mixed with the melted agarose and then poured into a Mini Gel Kit (CBS Scientific). Samples were mixed with loading buffer (30% glycerol, 0.25% bromophenol blue). Gels were run in 1x TBE-buffer. DNA was visualised with a UV lamp at 254 nm.
DNA extraction from Agarose Gels

a) Electroelution

The desired DNA fragment was excised from the agarose gel and put into a dialysis tube (SpectraPor 3, 3500 MWCO) with a little 1x TBE buffer. The bag was placed into the agarose-gel tank (in 1x TBE) and electrophoresis was performed for about half an hour at 80 volts (BioRad 500/200 power supply). The DNA eluted from the gel slice into the buffer and stuck to the membrane wall. To release the fragment, the poles were swapped and the electrophoresis run in the opposite direction for 15 seconds at 100 volts. The liquid was pipetted into an Eppendorf tube and precipitated with NaOAc and ethanol (see precipitation of DNA). The supernatant was discarded and the pellet washed once with 100 µl ice-cold 70% ethanol and dried. The DNA was dissolved in 20 µl sterile nanopure water or sterile 1x TE buffer.

b) QIAquick gel extraction

The protocol for the QIAquick Gel Extraction Kit (QUIAGEN) was used. The DNA fragment was excised from the agarose gel with a clean, sharp scalpel. The gel slice was kept as small as possible and weighed in a colorless tube. 3 volumes of Buffer QG to 1 volume of gel (100 mg ~ 100 µl) were added. The tube was then incubated at 50°C for 10 minutes or until the gel slice had completely dissolved. To help dissolution, the tube was vortexed every 2-3 minutes during the incubation. After the gel slice had dissolved completely, it was checked that the colour of the mixture was yellow. If the
colour of the mixture was orange or violet, 10 μl of 3 M sodium acetate, pH 5.0, was added to turn the colour yellow. To increase the yield of DNA fragments (<500 bp and >4 kbp), 1 gel volume of isopropanol was added. A QIAquick spin column was placed in one of the 2 ml collection tubes provided. To bind the DNA, the sample was applied to the QIAquick column and centrifuged for 1 minute. The flow-through was discarded and the column placed back in the same collection tube. The column was then washed by adding 0.75 ml of PE Buffer and centrifuging for 1 minute. Again the flow-through was discarded and the QIAquick column centrifuged for an additional 1 minute. The QIAquick column was placed into a clean 1.5 ml microfuge tube (cut off lid). To elute the DNA, 30 μl of Buffer EB (10 mM Tris-Cl, pH 8.5) was added to the center of the column and the tube was left to stand for 1 minute. It was centrifuged for 1 minute at maximum speed. The DNA was stored at -20°C.

Precipitation of DNA

DNA was precipitated by adding 1/10 vol. of 3 M NaOAc, pH 5.0, and 2-3 vol. of ethanol. The tube was placed in the freezer for at least 30 minutes. The precipitate was centrifuged for 10 minutes at 4°C. The supernatant was pipetted off. The dried DNA pellet was dissolved in 20 μl of sterile 1x TE buffer and stored at -20°C.
In vitro mutagenesis by the Kunkel method

The Muta-Gene® Phagemid In Vitro Mutagenesis Kit Version 2 (BIO-RAD) was used to introduce point mutations. This kit is based on a method described by Kunkel (Kunkel, 1985, Kunkel et al., 1987) which provides a selection against the non-mutagenised strand of a double-stranded DNA. When DNA is synthesised in a dut ung double mutant bacterium, the nascent DNA carries a number of uracils substituted for thymine as a result of the dut mutation, which inactivates the enzyme dUTPase and results in high intracellular levels of dUTP. The ung mutation inactivates uracil N-glycosylase, which allows the incorporated uracil to remain in the DNA. This uracil-containing strand is then used as the template for the in vitro synthesis of a complementary strand primed by an oligonucleotide containing the desired mutation. When the resulting double-stranded DNA is transformed into a cell with a proficient uracil N-glycosylase, the uracil-containing strand is inactivated with high efficiency, leaving the non-uracil-containing survivor to replicate. Typical mutagenesis frequencies obtained with the Muta-Gene phagemid kit are greater than 50%.

CJ236 cells containing the pBSE plasmid with the desired insert was streaked out onto a LB plate containing chloramphenicol and grown at 37°C until distinct colonies appear (usually overnight). An isolated colony was picked and placed in 5 ml of LB containing 50 μg/ml ampicillin and 30 μg/ml chloramphenicol. The culture was incubated with shaking at 37°C overnight. Then 50 ml of 2xYT media containing ampicillin and chloramphenicol was inoculated with 1 ml of the overnight culture of CJ236 containing phagemid, incubated at 37°C with shaking to an OD₆₆₀ of 0.3, which took from 1 to 4
hours. This corresponds to approximately $1 \times 10^7$ cfu/ml. M13KO7 helper phage ($1 \times 10^{11}$ pfu/ml) was added to obtain a MOI (multiplicity of infection) of around 20. After incubation for 1 hour, kanamycin (50 mg/ml) was added to a final concentration of 70 μg/ml. Incubation was continued overnight.

The overnight culture was held at 65°C for 15 minutes, then 30 ml of the culture were transferred to a 50 ml Oak Ridge centrifuge tube and centrifuged at 12'000 rpm (Sorvall RC-5B, SS34 rotor) for 15 min. The supernatant was recentrifuged in a fresh tube at 4°C. The second supernatant, transferred to a 50 ml falcon tube, was incubated at RT for 30 min with 150 μg RNase A. To the phagemid-containing supernatant, 1/4 volume 3.5 M ammonium acetate/20% PEG-6000 was added, mixed well and placed on ice for at least 30 minutes (preferably in a coldroom). The phagemids were collected by centrifuging at 12'000 rpm in an Oak Ridge tube for 15 min. The supernatant was decanted and discarded. The pellet was drained well, then dissolved in 200 μl of high salt buffer (300 mM NaCl, 100 mM Tris pH 8.0, 1 mM EDTA). The resuspended phagemids were stored in a 1.5 ml microcentrifuge tube at 4°C.

The entire phagemid stock was extracted twice with an equal volume of phenol/chloroform (phenol:chloroform:isoamyl alcohol, 49.5:49.5:1), and several times with chloroform/isoamylalcohol (24:1 v/v) until the layer between the phases disappeared. The DNA was precipitated by adding 1/10 vol. of 3 M NaOAc pH 5.0 and 2.5 vol. of ethanol and keeping the tube at minus 70°C for at least 30 minutes. It was centrifuged for 15 min in the cold, the supernatant removed and the pellet washed gently with 80% ice-cold ethanol, dried and dissolved in 20 μl 1x TE.
To anneal the phosphorylated primer and uracil-containing DNA, mutagenic oligonucleotide, 10x annealing buffer (1 μl) and water to 10 μl total volume were added to a microcentrifuge tube, heated to 70°C in a waterbath, and allowed to cool slowly. After that, the reaction mix was placed on ice. 10x synthesis buffer (1 μl), T4 DNA Ligase (1 μl, 3 units) and T7 DNA polymerase (1 μl, 0.5 units) were added to the cold solution and kept on ice for 5 min. The reaction was incubated at 25°C for 5 minutes and finally at 37°C for 90 minutes. 90 μl of TE stop buffer (10 mM Tris pH 8.0, 10 mM EDTA) were added to the reaction. The reaction products were transformed into E.coli DH5α competent cells and analysed for the desired mutation.

**Polymerase Chain Reaction (PCR)**

Primers used for PCR were purchased from Microsynth (Switzerland). All PCR's were performed on a Biometra UNO Thermocycler.

In a PCR tube (0.5 ml), 0.2 ng template (i.e. 1 μl of 1:100 dilution of a Wizard DNA Miniprep), 100 ng of each primer (~16 pmol), 2 μl dNTP mix (12.5 mM), 5 μl 10x Taq Pol buffer, 2 μl 50 mM MgCl₂, and sterile deionised water to 50 μl were mixed. The program was started and paused at 90°C, then the reaction mix was placed into the thermoblock, Taq Polymerase was added and the reaction topped with 30 μl of mineral oil. The program was continued. 29 cycles were performed. The reaction was denatured for 1 minute at 94°C. The primers were annealed at temperatures between 47°C and 55°C for 1-2 min. Chain extension took place for 1 min at 72°C.
Expression and of MASH-bHLH, MM17 and E12

All proteins were produced in BL21(DE3) cells containing the pLysS plasmid (Studier, 1991) from the T7 promoter in the plasmid pJGetita, a derivative of pET3a (Studier & Moffat, 1986). The MASH-1 and E12 cDNAs were from rat (Johnson et al., 1990; Lo et al., 1991).

Single colonies were picked from an agar plate and grown overnight in 5 ml of 2xYT medium. 4 ml of the cultures, 400 µl of ampicillin solution (50 mg/ml) and 400 µl of chloramphenicol solution (34 mg/ml ethanol) were added to 0.5 l 2x YT medium. Cells were grown at 37°C with shaking until the OD₆₀₀ of 0.4 - 0.5 (Varian DMS 300 UV/VIS Spectrophotometer) was reached. Then 400 µl of 1 M IPTG solution were added. Prior to adding the IPTG, 1 ml of the culture was extracted to use for gel analysis. Three more 1 ml samples were removed after 60, 120 and 180 minutes. Cells were centrifuged in 500 ml plastic bottles for 15 min at 5000 rpm and 4°C (Du Pont Instruments Sorvall RC-5C plus Refrigerated Superspeed centrifuge). The supernatant was discarded, and the pellet was stored at -20°C. Each 1 ml sample was centrifuged for 20 seconds in an Eppendorf centrifuge and the supernatants were discarded. 50 µl of 5x Sample-buffer and 5 µl of 1 M DTT were added to the pellet. The samples were vortexed and sonicated. After heating 2 min at 95°C, the samples were loaded and run on a SDS-PAGE gel.
**Protein Purification**

The cells were lysed by thawing in the presence of 3 ml of water/g wet cells and 1 mM PMSF. The cells were resuspended by vortexing or thorough shaking, then 2 volumes of lysis buffer (100 mM ammonium acetate, 100 mM sodium chloride, and 100 mM 2-mercaptoethanol) were added. The suspension was sonicated (Sonicator W-37, Microtip, Heat Systems Ultrasonics Inc.) for 10 min at room temperature and then centrifuged (Sorvall SS34) at 15000 rpm for 15 min. The pellet was resuspended in 20 ml of lysis buffer and recentrifuged. This process was repeated once, then the pellet was washed twice with 20 ml of 5 mM sodium acetate (pH 5.0) and centrifuged. The pellet was subsequently dissolved in urea buffer (8 M urea, 5 mM sodium acetate, pH 5.0, 100 mM 2-mercaptoethanol). To help dissolving the pellet, the sample was sonicated for 10 min at room temperature, then centrifuged (Sorvall SS34) at 15000 rpm for 15 min. The solution was applied to a column containing 30 ml of Bio-Gel CM A ion-exchange resin (Bio Rad), that was previously washed with 40 ml of 0.1 M sodium hydroxide, then rinsed with water to neutrality and equilibrated with urea buffer. The column was washed extensively with urea buffer and the protein eluted with 2-3 column volumes of urea buffer containing 1 M sodium chloride. Fractions containing the desired protein were combined and dialysed overnight against 20 volumes of urea buffer. The protein was further purified by preparative HPLC on a Resource-S sulfonate ion-exchange column (Pharmacia) using a linear gradient from urea buffer to urea buffer containing 1 M sodium chloride. The collected fractions were pooled and concentrated by ultrafiltration (Amicon,
YM3 membrane), and the buffer was exchanged to 5 mM sodium acetate pH 5.0 and 5 mM DTT. The protein was stored at -20°C.

**Preparation of Polyacrylamide-Gels**

The Running Gel was prepared by adding to a graduated cylinder solution A, 12.5 ml of solution B, 30 ml of solution J, and bringing the volume to 100 ml with ddH₂O. The solution was filtered and degassed. Then 500 μl of solution C, 30 μl of TEMED, and 500 μl of 10% ammonium persulfate were added. The solutions were poured into the assembled gel plates and then the solutions were covered with a thin layer of water saturated butanol. The spacers were inserted. When the gel polymerised completely (approx. 1 hour), the Stacking Gel (5%) was prepared by adding 8.3 ml solution A, 6.25 ml solution D, 2.5 ml solution J, and bringing the volume of the solution up to 50 ml with ddH₂O. The solution was filtered and degassed, 250 μl of solution C, 50 μl of TEMED, and 500 μl of 10% ammonium persulfate were added. The gels were rinsed with nanopure water to remove the butanol layer and the Stacking Gel solution was poured into the gel plates, the spacers inserted and the solution was left to polymerise. Once the solution polymerised, the gels were removed from the gel holder, wrapped in wet paper towel and cling film and stored at 4°C.
Solution A (30% acrylamide): 60 g acrylamide
1.6 g N,N'-methylen-bis-acrylamide
Water to 200 ml

Solution B (1.5 M Tris-HCl pH 8.8): 18.17 g Tris base dissolved in 80 ml water
Adjust pH to 8.8 with HCl
and add water to 100 ml

Solution C (20% SDS): 20 g SDS, add water to 100 ml

Solution D (0.5 M Tris-PO₄ pH 6.8): 6g Tris base, dissolve in 80 ml of water
Adjust pH with H₃PO₄
and add water to 100 ml

Solution J (50% Glycerol): 50 ml Glycerol (99%) and 50 ml of water
<table>
<thead>
<tr>
<th>Solution</th>
<th>Solution</th>
<th>Solution</th>
<th>Solution</th>
<th>Solution</th>
<th>1.5 M</th>
<th>20% SDS</th>
<th>0.5 M</th>
<th>50% Glycerol</th>
<th>TEMED</th>
<th>H₂O</th>
<th>10% A.P.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>J</td>
<td>Tris-Cl</td>
<td>Tris-PO₄</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30% acrylamide</td>
<td>1.5 M</td>
<td>20% SDS</td>
<td>0.5 M Tris-PO₄</td>
<td>50% Glycerol</td>
<td>Tris-Cl</td>
<td>Tris-PO₄</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Running gels</td>
<td>(%gel desired/3) in ml</td>
<td>1.25 ml</td>
<td>50 μl</td>
<td>(%gel desired/5) in ml</td>
<td>3 μl</td>
<td>to 10 ml</td>
<td>50 μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stacking gels</td>
<td>(%gel desired/3) in ml</td>
<td>50 μl</td>
<td>1.25 ml</td>
<td>(%gel desired/5) in ml</td>
<td>10 μl</td>
<td>to 10 ml</td>
<td>100 μl</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Measurement and Calculation of Protein Concentration

Measurements were made at wavelengths of 210 nm, 215 nm and 220 nm using a Varian DMS 300 UV/VIS Spectrophotometer. A clean quartz cuvette (HELMA) was used to hold the samples. The measurements were carried out in the following order and were repeated three times:

1. 1 ml ddH$_2$O  
2. + 5 μl Reference buffer  
The cell was carefully washed with ddH$_2$O  
3. 1 ml ddH$_2$O  
4. + 5 μl Protein solution  

Protein concentration was calculated from the following equation (Wetlaufer et al., 1962):

$$c \text{ [mg/ml]} = \frac{[(\text{Pr-H}_2\text{O})-(\text{Re-H}_2\text{O}) \times 201]}{X_\lambda}$$

Where

- Pr = Protein in ddH$_2$O  
- Re = Reference buffer in ddH$_2$O  

$$X_{210} = 20$$  
$$X_{215} = 15$$  
$$X_{220} = 11$$

The concentrations were calculated at each wavelength in every run and the mean value of all the results was taken to be the calculated concentration.
**Mass spectroscopy**

MALDI-TOF (Matrix Assisted Laser Desorption Ionisation-Time Of Flight) mass spectra were obtained by Mr. H. Hediger (ETH Service), by Dr. P. James at the ETH Zurich or at the mass spectrometry service at the University of Birmingham. Sinapinic acid was used as a matrix.

**CD-Spectroscopy**

Circular dichroism (CD) spectra were recorded on a Jasco J600 or Jasco J710 spectrometer at 25°C. Spectra were analysed using the software supplied with the spectrometer. Cuvettes (HELMA) with a pathlength of 10 mm were used. The resolution was 0.2 nm, the bandwidth 1.0 nm. Buffers for the measurements are indicated.

**Electrophoretic mobility shift assay (EMSA)**

Oligonucleotides used for the EMSA experiments were synthesised by Microsynth (Balgach Switzerland) or the Institute of Zoology at the University of Zurich.

Oligonucleotides were labelled by mixing 2\mu l 10x Kinase Buffer (0.7 M Tris-HCl pH 7.6, 100 mM MgCl₂, 50 mM DTT), 2 \mu l 100 mM MgCl₂, 2 \mu l oligo 1 (0.5 \mu g), 10 \mu l ddH₂O, 1 \mu l (\gamma-³²P)-ATP (Amersham, 250 \mu Ci) and 1 \mu l of polynucleotide kinase. The solution was incubated for 2 hours at 37°C. The
second oligo was added, the solution heated to 90°C and left to cool to room temperature. The mixture was diluted with 480 μl of water to a final concentration of 1 ng/μl double-stranded oligonucleotide.

Aliquots of proteins were serially diluted into EMSA buffer (50 mM Tris-HCl pH 7.9, 6 mM MgCl₂, 40 mM ammonium sulphate, 0.2 mM EDTA, 1 mM DTT, 5% glycerol). This solution was incubated in the presence of 10 nM labelled oligonucleotide, 2.5 μl of 25% CHAPS, to avoid formation of higher aggregates, and water to 20 μl for 10 min at room temperature. Samples were applied to 4% polyacrylamide and gels run in 0.9x TAE pH 7.9 at 70 volts at 4°C in a vertical gel electrophoresis apparatus (V15-17, BRL Life technologies, Inc.). After electrophoresis, the gels were dried and exposed to Kodak X-OMAT-S film overnight at -70°C. Quantitative data were obtained with a Instant Imager (Packard) or a Phosphor Imager (Molecular Dynamics) using system software.

The fraction Θ 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 from EMSA titrations. Plotting Θ against the concentration of unbound protein [P] allowed the determination of the concentration [P]₁/₂, where half of the protein binding sites were occupied. The concentration of free protein could be calculated as [P] = [P]₁₋₂[P₂D], where [P]₁ and [P₂D] are the concentrations of the total amount of protein added and of the complex, respectively (Meierhans et al., 1995).

Dissociation constants were calculated as Kᵤ = ([P]₁/₂)².
**Gels for Band Shift Assays**

Gels was prepared by filtering 17 ml of a 30% acrylamide solution, 2.6 ml of 50 x TAE buffer and 110 ml nanopure water into a vacuum flask. The solution was degassed for at least 15 minutes. 2.6 ml CHAPS (25% w/v), 125 µl TEMED and 500 µl 10% ammonium sulfate solution were added. The mixture was poured immediately between the assembled glass-plates (front plate: 16 x 19.7 x 0.3 mm; back plate: 19 x 19.7 x 0.3 mm; spacer: 1 mm) fixed with two clamps on each side. The comb was inserted and the plates were kept horizontally at RT until the gel polymerised, then stored at 4°C.

**Western Blots**

The nitro-cellulose membrane was equilibrated for 1 hour in transfer solution. After the samples were run on a 15% SDS-PAGE Gel, the proteins were transferred onto a nitro-cellulose membrane for 2.5 hours at 250 mA at 4°C. The blot was stained with Red Ponceau to locate and mark protein bands with a blue pen and rinsed afterwards with PBS. Then the blot was blocked for 30 minutes in Blotto saturation Solution at 37°C. After washing in Blotto Washing solution, the membrane was incubated for 1 hour at 37°C with primary antibody (Rabbit anti-MASH-bHLH at 1:300 dilution). The membrane was washed three times with Blotto Washing solution and incubated with a secondary antibody (Goat anti-Rabbit/Alkaline Phosphatase Conjugate, 1:70'000 dilution in Blotto Washing Solution). The blot was rinsed with PBS and developed by addition of Nitroblue Tetrazolium and 5-bromo-4-chloro-3-indolyl-phosphate to the alkaline phosphatase buffer.
10x Transfer Buffer pH 8.5:
- 29g Glycine
- 58g Tris base
- 800ml of ddH₂O
- adjust pH with HCl
- add ddH₂O to 1000ml

Transfer solution:
- 200ml 10x Transfer Buffer pH 8.5
- 7.4ml 10% SDS
- 400ml Methanol
- add ddH₂O to 2000ml

Red Ponceau:
- 0.2% w/v Ponceau S; 200mg/100ml
- 1% acetic acid; 1ml/100ml

10x PBS:
- 80g NaCl
- 2g KCl
- 14.4g Na₂HPO₄
- 2.4g KH₂PO₄
- in 800ml of Water
- adjust to pH 7.4 with HCl
- add ddH₂O to 1000ml / autoclave

10x Blotto:
- 1M Tris pH 8.0 (60.57g)
- 1.5M NaCl (43.83g)
- add water to 500ml

Blotto Saturation Solution:
- 1x Blotto 2ml 10x Blotto
- 5% Dry Milk 1g Dry Milk
- 0.05% Tween 20 10μl Tween 20
- to 20ml with ddH₂O
Isothermal Titration Calorimetry

ITC experiments were carried out on a Omega (for MASH-BHLH) and a MCS (for MM17) isothermal titration calorimeter (MicroCal, Inc., Northampton, MA), consisting of a measuring and computing unit, a waterbath and computer. The system was described by Wiseman and co-workers (Wiseman et al., 1989).

Oligonucleotides were purchased from Integrated DNA Technologies.

Protein and DNA solutions were dialysed in the same buffer to avoid even minimal differences that could have an effect on the ITC measurement. They were dialysed 3 times against 400 ml and one time against 800 ml buffer, which was kept and used in the experiments. Protein and DNA concentrations were then determined.

The waterbath that cools the ITC system was set 5 degrees lower than the reaction temperature at least 12 hours before the first measurement to equilibrate the cell. 2 ml of a 5 μM (MASH-BHLH) or 1 μM (MM17) DNA solution were prepared. All samples were degassed for 10 min. In the meantime, the cell was rinsed three times with 1.5 ml of ice-cold buffer and drained well. The DNA solution was injected into the cell. The syringe (with stirrer blade) was loaded with the protein solution (184 μM MASH-BHLH or 50 μM MM17) and mounted onto the cell. The experiments were directed by the Microcal software from a connected computer. 1 μl and 12x 8 μl in the case of MASH-BHLH or 1 μl and 18 to 21 x 4 μl in the case of MM17 protein solution were injected. Time intervals between the injections were 240 seconds.

Analysis was carried out with the supplied software (Microcal Data Analysis)
that gave directly the evolved heat of the reaction, $\Delta H$, the stoichiometry and the Association constant $K_A$.

Buffers for ITC with MASH-BHLH:
- Low-salt MES buffer: 10 mM MES-KOH pH 6.0
- High-salt MES buffer: 10 mM MES-KOH pH 6.0, 100 mM KCl
- PIPES-6 buffer: 10 mM PIPES pH 6.0
- PIPES-7 buffer: 10 mM PIPES pH 7.0

Buffers for ITC with MM17:
- Low-salt MES buffer: 10 mM MES-KOH pH 6.0, 5 mM KCl
- Low-salt PIPES buffer: 10 mM PIPES-KOH pH 7.0, 5 mM KCl
- 25P7 buffer: 10 mM PIPES-KOH pH 7.0, 25 mM KCl
- Low-salt Tris buffer: 10 mM Tris-HCl pH 8.0, 5 mM KCl

**Computer Graphics**

Pictures were created on Silicon Graphics Computers (O2) using the programs Molscript, Version 1.4 (Kraulis) or VMD (Visual Molecular Graphics) (Humphrey et al., 1996). All presentations were rendered with Raster 3D (Bacon & Anderson, 1988; Meritt & Murphy, 1994). Alterations in the molecular structure from PDB data files were done with MacMOMO, Version 2.0.4 (Dobler, ETH Zurich), or Sybyl (Tripos Associates, Inc.).
References


Dobler, M.; Laboratorium für Organische Chemie, ETH Zürich


Olson, E.N. (1993) Circ. Res. 72, 1-6


Spolar, R.S. & Record, M.T., Jr. (1994) Science 263, 777-784


Curriculum Vitae

1976 - 1983 Primarschule in Glarus und Mitlödi
1983 - 1989 Besuch der Kantonsschule in Glarus
1989 Matura Typus B
1989 Arbeit auf der Post in Glarus
1990 Gebirgsrekrutenschule in Chur, dann Unteroffiziersschule Chur und Abverdienen
Oct. 1990 - 1995 Chemiestudium an der ETH in Zürich und Diplomarbeit bei Dr. Rudolf K. Allemann
1995 - 1998 Assistent am Laboratorium für Organische Chemie
1995 - 2000 Doktorarbeit bei Dr. Rudolf K. Allemann an der ETH Zürich und an der Universität Birmingham (UK).