Basic−Helix−Loop−Helix Region of Tal: Evaluation of Structure and DNA Affinity

Patricia Bishop, Indraneel Ghosh, Cory Jones, and Jean Chmielewski*

Department of Chemistry, Purdue University West Lafayette, Indiana 47907

Received January 31, 1995

The product of the protooncogene *TAL1* has been identified as a putative member of the basic−helix−loop−helix class of transcription factors due to high sequence homology with other bHLH proteins of known structure,[1] such as immunoglobulin enhancer binding protein E47 (E47),[2] USF,[3] and Max[4] (Figure 1). Rearrangements and translocations within the *TAL1* gene have been found in 30% of patients with T cell acute lymphoblastic leukemia (T-ALL); approximately 25% of the patients exhibited a 90 kb deletion in the *TAL1* gene,[5] and in an additional 3% of T-ALL patients the *TAL1* gene had undergone a translocation from its normal location on chromosome 1 into the T-cell receptor δ chain gene on chromosome 14.[6] The *TAL* gene is expressed in early hematopoietic tissues, leading to speculation that the Tal protein plays a role in differentiation or commitment events during hematopoiesis.[7] Although the precise function and DNA binding of Tal have yet to be demonstrated, studies have determined that Tal functions similarly to tissue-specific bHLH proteins involved in transcriptional regulation by forming DNA-binding heterodimers with E47 and E12, two ubiquitously expressed HLH proteins.[8,9] However, unlike the tissue-specific bHLH proteins, homodimeric complexes of Tal have not been detected.[9]

Since dimerization and DNA binding are generally constant features of the bHLH motif, it is interesting that Tal does neither on its own. Consequently, we have prepared the bHLH region of Tal (**1**) to test if homodimers may be formed with this smaller segment. If homodimers are obtained with **1**, then DNA binding may be assayed.

Protein **1** was synthesized by a solid phase synthesis approach using an Fmoc-based stragegy[10] on the Wang resin.[11] Reverse phase HPLC was used to purify the peptide to homogeneity to give **1** in an overall yield of 10%.[12] This protein was judged to be greater than 95% pure by analytical HPLC and SDS-PAGE, and the sequence was confirmed by electrospray mass spectrometry, amino acid analysis, and amino-terminal sequencing.

```
        Basic      Helix I            Loop              Helix II
E47   ERRMANNARERVRVRDINEAFRELGRMCQMHLKSDKAQT-----KLLILQQAVQVILGLEQQVR (2)
Tal   VRRIFTNSRERWRQQNVNGAFAELRKLIP----THPPDKKLS--KNEILRLAMKYINFLAKLLN (1)
Max   DRRAHHNALERKRRNELKRSFHSLRDHVP----SLEQGEKAS--RAQILDKATEYIQTMRRKND
USF   KRRAQHNEVERRRRDKINNWIVQLSKIIP----DCSMESTSGQSKGGILSKACDYIQELRQSNH
```

Figure 1. Sequence alignment of proteins containing the basic−helix−loop−helix motif.

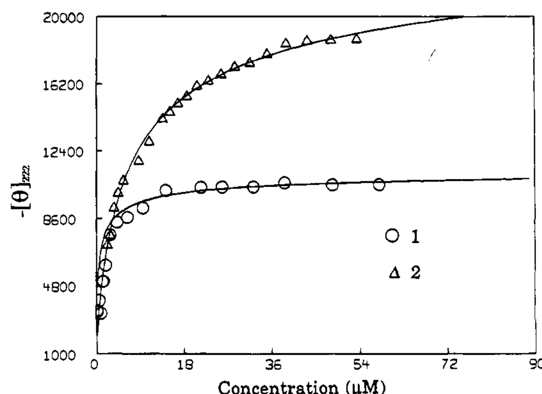

Figure 2. Binding isotherms for **1** and **2** in 10 mM phosphate, 100 mM NaCl, 1 mM DTT, pH 6.5 buffer at 24 °C fitted with a monomer−dimer equilibrium scheme.

Size exclusion chromatography of **1** (150 μM in 10 mM phosphate, 150 mM NaCl, 1 mM DTT, pH 6.0 buffer) on Sephadex G-50 as compared to protein standards provided an apparent molecular weight of 15 560, which corresponds to an aggregation state of 2.2 for **1**. Sedimentation equilibrium according to the method of Pollet[13] (150 μM **1**) provided further evidence for dimerization; a molecular weight of 14 700 was calculated, which corresponds to an aggregation state of 2.1 for **1**. Cross-linked complexes of **1** were also obtained with the bifunctional cross-linking reagent Sulfo-EGS (Pierce). At 100 μM, 20% of **1** was cross-linked to an upper band with an approximate molecular weight of 14 000, which corresponds to a cross-linked dimer.

Dimeric **1** was also helical (32% at 95 μM in 10 mM phosphate, 100 mM NaCl, 1 mM DTT, pH 6.5 buffer at 24 °C) as judged by circular dichroism. As the concentration of **1** was lowered, however, there was a concomitant decrease in the helical content of **1** as it dissociated from the aggregated to the monomeric state. This concentration dependence was analyzed according to various monomer−*n*-mer equilibria using a nonlinear regression program (MLAB), and a best fit was obtained where *n* (the aggregation state) was equal to 2 (Figure 2).[16] In this way a dissociation constant (K_d) of 2.1 ± 0.1 μM was obtained for **1**, which is similar to the K_d obtained for the bHLH region of IEB E47 (**2**) using the same method (6.8 ± 2.2 μM).[17]

With the knowledge that the bHLH region of Tal does dimerize, a gel mobility shift assay was used to assess the affinity of **1** for a number of DNA probes containing transcriptional enhancer (E box) sequences (CAGGTG (κE2), CAGCTG,

* Author to whom correspondence should be addressed.

(1) Visvader, J.; Begley, C. G. *Trends Biochem. Sci.* **1991**, *16*, 330−333.
(2) Ellenberger, T.; Fass, D.; Arnaud, M.; Harrison, S. C. *Genes Dev.* **1994**, *8*, 970−980.
(3) Ferre-D'Amare, A. R.; Pognonec, P.; Roeder, R. G.; Burley, S. K. *EMBO J.* **1994**, *13*, 180−189.
(4) Ferre-D'Amare, A. R.; Prendergast, G. C.; Ziff, E. B.; Burley, S. K. *Nature* **1993**, *363*, 38−45.
(5) (a) Brown, L.; Cheng, J. T.; Chen, Q.; Siciliano, M. J.; Crist, W.; Buchanan, G.; Baer, R. *EMBO J.* **1990**, 3343−3351. (b) Bernard, O.; Lecointe, N.; Jonveaux, P.; Souyri, M.; Mauchauffe, M.; Berger, R.; Larsen, C. J.; Mathieu-Mahul, D. *Oncogene* **1991**, *6*, 1477−1488.
(6) (a) Chen, Q.; Cheng, J. T.; Tsai, L. H.; Schneider, N.; Buchanan, G.; Carroll, A.; Crist, W.; Ozanne, B.; Sicilliano, M. J.; Baer, R. *EMBO J.* **1990**, *9*, 415−424. (b) Aplan, P. D.; Lombardi, D. P.; Ginsberg, A. M.; Cossman, J.; Bertness, V.; Kirsch, I. R. *Science* **1990**, *250*, 1426−1429.
(7) Begley, G. C.; Aplan, P. D.; Dennings, S. M.; Haynes, B. F.; Waldmann, T. A.; Kirsch, I. R. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 10128−10133.
(8) (a) Hsu, H. L.; Cheng, J. T.; Chen, Q.; Baer, R. *Mol. Cell. Biol.* **1991**, *11*, 3037−3042. (b) Hsu, H. L.; Huang, L.; Tsan, J. T.; Funk, W.; Wright, W. E.; Hu, J. S.; Kingston, R. E.; Baer, R. *Mol. Cell. Biol.* **1994**, *14*, 1256−1265.
(9) (a) Hsu, H. L.; Wadman, I.; Baer, R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3181−3185. (b) Hsu, H. L.; Wadman, I.; Tsan, J. T.; Baer, R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5947−5956.
(10) Atherton, E.; Sheppard, R. C. In *The Peptides*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, 1987; Vol. 9, pp 1−38.
(11) Wang, S. S. *J. Am. Chem. Soc.* **1973**, *95*, 1328−1333.

(12) HPLC conditions: Waters Delta-Pak C_{18} radial compression column (5 cm × 28 cm), with a linear gradient of 25−60% CH_3CN/H_2O (0.1% TFA) over 60 min.

(13) Pollet, R. J. *Methods Enzymol.* **1985**, *117*, 3−27. Pollet, R. J.; Haase, B. A.; Standaert, M. L. *J. Biol. Chem.* **1979**, *254*, 30−33.
(14) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108−4116.
(15) Lyu, P. C.; Sherman, J. C.; Chen, A.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5317−5320.
(16) Fairman, R.; Beran-Steed, R. K.; Anthony-Cahill, S. J.; Lear, J. D.; Stafford, W. F., III; DeGrado, W. F.; Benfield, P. A.; Brenner, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10429−10433.

(17) Fairman *et al.* report a K_d for **2** of 4.5 μM, which is in agreement with our value.[16]

0002-7863/95/1517-8283$09.00/0 © 1995 American Chemical Society

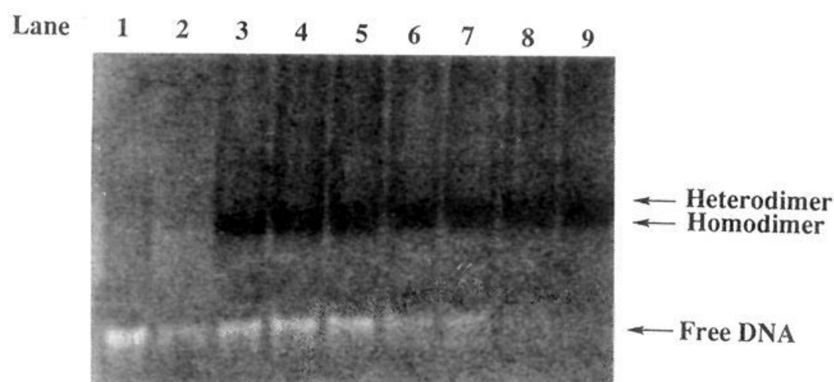

Figure 3. Gel mobility shift assay: lane 1, DNA alone (7.5 μM); lane 2, DNA and **1** (8 μM); lane 3, DNA and **2** (8 μM); lane 4, DNA and **2** (8 μM) and **1** (4 μM); lane 5, DNA and **2** (8 μM) and **1** (7.5 μM); lane 6, DNA and **2** (8 μM) and **1** (11 μM); lane 7, DNA and **2** (8 μM) and **1** (14.5 μM); lane 8, DNA and **2** (8 μM) and **1** (18 μM); lane 9, DNA and **2** (8 μM) and **1** (21.5 μM). Buffer: 1 mM phosphate, 10 mM NaCl, 0.1 mM DTT, 0.05 mM EDTA, 10% glycerol, and 1 μg of poly-dIdC per experiment. All bands were silver stained.

CAACTG, and CAGATG). The absence of a shifted band in these experiments indicated that **1**, although it does homodimerize, does not have specific affinity for the DNA sequences used even at protein concentrations approaching 100 μM. To determine if **1** existed in the same conformation as the bHLH region of full-length Tal, DNA binding of heteodimeric complexes of **1** and the bHLH region of E47 (**2**) was investigated with a DNA probe containing the E box sequence CAGATG, a sequence for which full-length Tal and IEB E47 had previously shown affinity (Figure 3).[8a] Beginning with a concentration of **2** (8 μM), which bound approximately half of the DNA probe, addition of **1** (4 μM) gave rise to a second shifted band. At higher concentrations of **1** ($\sim$20 μM) essentially all of the DNA probe had been converted to the second, higher shifted band. Western blot analysis of the DNA-binding gel with polyclonal antibodies raised against **1** confirmed the presence of **1** in this shifted band.

Unlike the dimerization experiments performed with full-length Tal (331 residues),[9] our experiments with the bHLH region of Tal provide evidence for dimer formation with a dissociation constant in the same range as the DNA-binding, bHLH region of IEB E47. Although stable dimers of **1** form, there is still no evidence for DNA binding as was also observed with full-length Tal. There are HLH proteins which self-associate but do not bind DNA, such as Id,[16] which contain helix-breaking residues in what would be the DNA-binding region. Protein **1**, however, is without proline residues in this region and also has all of the conserved basic residues normally associated with DNA binding.

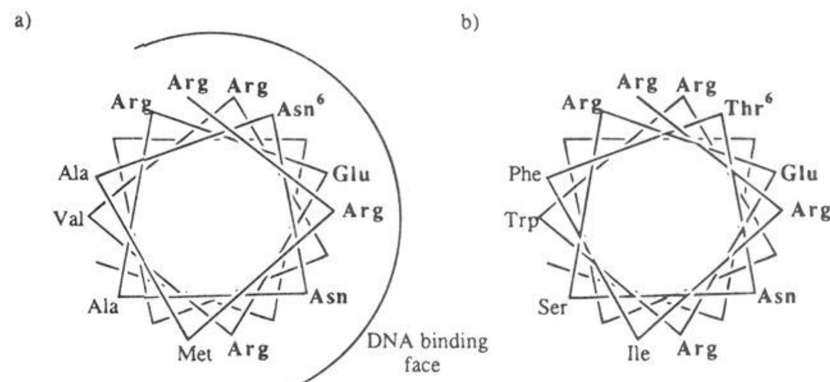

Figure 4. Helical wheel representations of the basic regions of (a) E47 and (b) Tal.

If one compares the helical basic region of E47 to Tal, the residues of E47 which interact with DNA[2] are highly conserved within Tal, with the exception of an Asn(6)−Thr(6) change (Figure 4). To test if this residue is responsible for the lack of DNA binding, we prepared an Asn(6) mutant of **1**. Although this protein dimerized, no DNA binding was detected. More complex structural issues, such as the orientation of the basic regions and steric interactions between the bHLH regions, may also play a role in the lack of DNA binding with homodimers of **1**.

It is interesting that although homodimers of **1** lack DNA affinity, heterodimers between **1** and the bHLH of E47 (**2**) do not. HLH proteins which lack DNA binding, such as Id, also form multimeric complexes with E47. In the case of Id, however, the DNA binding of E47 is inhibited, presumably because a single E47 basic region is not sufficient for complexation.[16] If the sequence and the orientation of the basic region of **1** are not appropriate for homodimeric DNA binding, how is it that heterodimerziation with E47 overcomes these obstacles? Fusion and mutant proteins are currently being prepared to address this question.

Acknowledgment. We gratefully acknowledge the financial support of the NIH (5 R29 GM 46936-02), the American Cancer Society (Purdue Cancer Center), and the Monsanto Company.

Supporting Information Available: Characterization, size exclusion, sedimentation equilibrium, and DNA-binding data for **1** (4 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA950319X