Volume 267, number 1, 139–141 FEBS 08573 July 1990

High mobility group proteins 1 and 2 bind preferentially to brominated poly(dG-dC)·poly(dG-dC) in the Z-DNA conformation but not to other types of Z-DNA

T. Christen, M. Bischoff, R. Hobi and C.C. Kuenzle

Institut für Pharmakologie und Biochemie, Universität Zürich-Irchel, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

Received 10 May 1990

Three proteins from bull testis, previously thought to be Z-DNA-binding proteins but recently found to recognize brominated poly(dG-dC)·poly(dG-dC) by criteria different from the Z-conformation, were partially sequenced. Of these, the 31 kDa protein was identified as a member of the high mobility group 2 protein family, and the 33 kDa protein as a member of the high mobility group 1 protein family. Both proteins had molecular weights approximately 30% higher than expected, indicating considerable posttranslational modification. In contrast, the 58 kDa protein remained unidentified for lack of any significant homology with known protein sequences.

Z-DNA, Z-DNA-binding protein; High mobility group protein

1. INTRODUCTION

We have previously reported on the isolation of three Z-DNA-binding proteins from bull testis with molecular masses of 31, 33 and 58 kDa [1]. Identification as Z-DNA-binding proteins relied on their selective binding to brominated poly(dG-dC)·poly(dG-dC), an accepted model for Z-DNA [2], as opposed to the unbrominated polynucleotide in B-form. We have reassessed the conformational binding specificity of these 3 proteins with more natural Z- and B-DNA ligands. These were composed of circular DNA molecules, in which a $d(CG)_7$ insert was induced to adopt the Z-conformation upon supercoiling [3]. With these ligands we found no preferential binding to Z-DNA [4] leading us to conclude that the 3 proteins discriminate brominated from unbrominated poly(dG-dC)·poly(dG-dC) on the basis of chemical properties different from the Z- and B-conformation, respectively.

In parallel with the binding studies we initiated structural work aimed at identifying the 3 proteins by partial amino acid sequencing and comparison with known proteins. Here we report the successful identification of the 31 kDa and 33 kDa testis proteins as members of the HMG 2 and HMG 1 (HMG = high mobility group) protein families, respectively. In contrast, the 58 kDa protein remained unidentified for lack of any significant homology to known protein sequences.

Correspondence address: C.C. Kuenzle, Institut fur Pharmakologie und Biochemie, Universität Zurich-Irchel, Winterthurerstr. 190, CH-8057 Zurich, Switzerland

2. MATERIALS AND METHODS

2.1. *Protein isolation*

The proteins of molecular masses 31, 33 and 58 kDa were isolated from mature bull testis by method II as described [1], except that fractions containing the 58 kDa protein were never pooled but processed independently. Fractions were monitored by SDS-polyacrylamide gel electrophoresis [5] and by filter binding tests using brominated and unbrominated poly(dG-dC)·poly(dG-dC) as Z- and B-DNA ligands, respectively [1]. The purity of the 58 kDa protein was further checked by reverse-phase HPLC [1]. The final preparations contained less than 5% impurities and bound brominated poly(dG-dC)·poly(dG-dC) with the same dissociation constants as determined previously [1].

2.2. *Preparation of the purified proteins for amino acid sequencing*

The 31 kDa and 33 kDa proteins were directly subjected to amino acid sequencing. Because, during the first three cycles of the Edman degradation, glycine signals were obscured by an unidentified contamination (probably protein-bound DNA), the 31 kDa protein was also sequenced following further purification by SDS-polyacrylamide gel electrophoresis [6] and transfer onto a PVDF membrane [7].

The 58 kDa protein could not be sequenced directly probably because of N-terminal blockage. Therefore, the protein was digested with the endoproteinase Asp-N (Boehringer, Mannheim). For this purpose, the protein solution (100 μg in 50 μl of 50 mM Tris buffer, pH 7.8) was supplemented with 10 μl acetonitrile (to effect unfolding of the protein) and was incubated with Asp-N (2 μg in 50 μl of distilled water) at 37°C for 16 h. The reaction mixture was divided into 4 aliquots, and the peptides were separated by reverse-phase HPLC on a Nucleosil 300 column (microbore, 2 × 200 mm, C8, 5 μm; Stagroma, Zürich) at a flow rate of 200 μl/min. Gradient elution was with buffers HPLC-A (0.1075 vol.% trifluoroacetic acid in water) and HPLC-B (0.085 vol.% trifluoroacetic acid and 70 vol.% acetonitrile in water) in 3 steps: step 1 = from 0% to 25% buffer HPLC-B the gradient rose by 2.5% buffer HPLC-B/min; step 2 = from 25% to 70% buffer HPLC-B the gradient rose by 1.0% buffer HPLC-B/min; step 3 = from 70% to 100% buffer HPLC-B the gradient rose by 6.0% buffer HPLC-B/min. Effluents were monitored by absorption at 220 nm (Fig. 1). Corresponding peptide peaks from

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

4 independent runs were combined, checked for purity by reverse-phase HPLC as above and subjected to amino acid sequence analysis.

2.3. *Amino acid sequencing*

Gas-phase sequencing was performed with an Applied Biosystems Sequenator type 470 A.

3. RESULTS

3.1. *31 kDa protein*

The amino acid sequence of the 31 kDa protein was determined for residues 1–38:

1	G	K	G	D	P	N	K	P	R	G	10
11	K	M	S	S	Y	A	F	F	V	Q	20
21	T	**X**	R	E	E	H	K	K	K	H	30
31	P	D	A	S	V	N	F	**A**			38

This sequence matched that of the bovine HMG 2 protein [8,9] with 2 exceptions (emphasized by bold face type): residue 22 could not be identified in the 31 kDa protein (marked X) but is serine in HMG 2, and residue 38 was alanine in the 31 kDa protein but is serine in HMG 2. An ambiguity exists with respect to residue 36, which was asparagine in the 31 kDa protein but has been identified only partially in HMG 2 as being either asparagine or aspartic acid. From this we conclude that the 31 kDa protein is a member of the HMG 2 protein family. The failure to identify residue 22 is suggested to result from some unidentified form of posttranslational modification. This interpretation is supported by the greater apparent molecular mass of the 31000 testis protein (31000 as determined by SDS-polyacrylamide gel electrophoresis) compared with that of the authentic calf thymus HMG 2 (22500 as determined by SDS-polyacrylamide gel electrophoresis [10], or 24000–25000 as deduced from the incompletely determined primary structure [8,9]).

3.2. *33 kDa protein*

The amino acid sequence of the 33 kDa protein was determined for residues 1–37:

1	(G)	K	(G)	D	P	K	K	P	R	G	10
11	K	M	S	S	Y	A	F	F	V	Q	20
21	T	**X**	R	E	E	H	K	K	K	H	30
31	P	D	A	S	V	N	F				37

Apart from two uncertainties at positions 1 and 3 (rendered in parentheses) and an unidentified residue at position 22 (marked X and emphasized by bold face type), this sequence corresponded to that of the bovine HMG 1 protein [8,9,11]. In HMG 1, residues 1 and 3 are glycine (as also suggested in this work), and residue 22 is either serine [8,9] or cysteine [11]. As with the 31 kDa protein (see above), the failure to identify residue 22 might be attributed to some unidentified form of posttranslational modification. This interpretation is supported by the greater apparent molecular mass of the 33000 testis protein (33000 as determined by SDS-polyacrylamide gel electrophoresis) compared with that of the authentic calf thymus HMG 1 (23400 as determined by SDS-polyacrylamide gel electrophoresis [10] or 24907 as deduced from the primary structure [8,9,11]).

3.3. *58 kDa protein*

The 58 kDa protein could not be sequenced directly probably as a result of N-terminal blockage. The protein was therefore cleaved with the endoproteinase Asp-N (which releases peptides beginning with Asp), and the resulting peptides were separated by reverse-phase HPLC. Four peaks, numbered 1–4, were selected for amino acid sequence analysis (Fig. 1). Sequencing revealed that peaks 1, 2 and 4 were uniform, while peak 3 was composed of two peptides. The major peptide

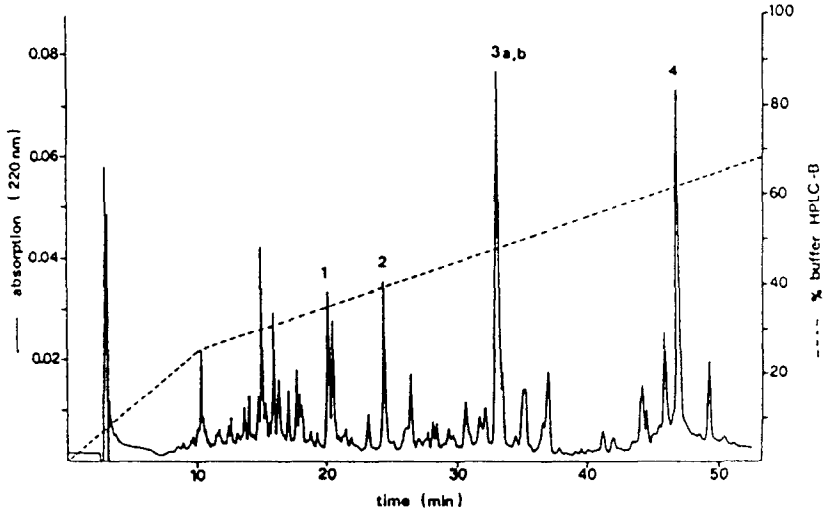

Fig. 1. Reverse-phase HPLC of peptide fragments derived from the 58 kDa protein by cleavage with the endoproteinase Asp-N. Peaks selected for amino acid sequence analysis are marked 1–4

was labeled 3a and the minor peptide 3b. The two peptides differed sufficiently in concentration to allow their simultaneous sequencing in a single run. The following sequences were obtained (uncertain residues are rendered in parentheses; unidentified residues are marked X):

Peptide 1
1 D P K G P P L L (E)E 10
11 D P G A K 15

Peptide 2
1 D P A K E E Q L T S 10
11 V L H L S G G G L G 20

Peptide 3a
1 D I V E E L K N P K 10
11 P P A L V M F Y A P 20

Peptide 3b
1 X N F X X T L K (R)K 10
11 K (H)Q P Q (V)P E T (T) 20

Peptide 4
1 D E G G S V Y H L S 10
11 D E D F D Q F V K E 20
21 H S S V L V M F (S)H 30

A computer search for sequence homologies was conducted with each of the 5 peptide sequences based on the data bases DNA-STAR, SWISS-PROT and MIPSX. However, in no instance did we find a significant homology. This indicates that the 58 kDa protein is an as yet unidentified testis protein.

4. DISCUSSION

Three proteins from bull testis, previously thought to be Z-DNA-binding proteins [1] but recently found to recognize brominated poly(dG-dC)·poly(dG-dC) by criteria different from the Z-conformation [4], were partially sequenced. Of these, the 31 kDa protein was identified as a member of the HMG 2 protein family, and the 33 kDa protein as a member of the HMG 1 protein family. Both proteins had molecular weights approximately 30% higher than expected, indicating considerable posttranslational modification. This might also explain their marked hydrophobicity [1], a feature not normally encountered with the HMG 1 and 2 proteins. In contrast, the 58 kDa protein remained unidentified for lack of any significant homology with protein sequences registered in 3 databases.

HMG proteins 1 and 2 have previously been tested for potential binding to Z-DNA with negative outcome [10]. This is consistent with our own results (this work and [4]). However, HMG 1 and 2 are known to have a preference for unusual DNA structures such as single-stranded DNA [12], cruciforms [10,13] and B–Z junctions [10]. It is possible that this selectivity is also reflected in the high-affinity binding to brominated poly(dG-dC)·poly(dG-dC).

Acknowledgements: We thank Dr P. Hunziker, Biochemisches Institut der Universitat Zurich, for amino acid sequencing. This work was supported by the Swiss National Science Foundation (Grants 31-8678.86 and 31-26117.89).

REFERENCES

[1] Gut, S.H., Bischoff, M., Hobi, R. and Kuenzle, C.C. (1987) Nucleic Acids Res. 15, 9691–9705.
[2] Moller, A., Nordheim, A. and Rich, A. (1984) Biochemistry 23, 54–62.
[3] Nordheim, A. and Meese, K. (1988) Nucleic Acids Res. 16, 21–37.
[4] Rohner, K.J., Hobi, R. and Kuenzle, C.C. (1990) submitted.
[5] Laemmli, U.K. (1970) Nature 227, 680–685.
[6] Moos, M., Nguyen, N.Y. and Liu, T.-Y. (1988) J. Biol. Chem. 263, 6005–6008.
[7] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[8] Walker, J.M., Gooderham, K., Hastings, J.R.B., Mayes, E. and Johns, E.W. (1980) FEBS Lett. 122, 264–270.
[9] Walker, J.M. (1982) in: The HMG Chromosomal Proteins (Johns, E.W. ed.) pp. 69–87, Academic Press, London.
[10] Hamada, H. and Bustin, M. (1985) Biochemistry 24, 1428–1433.
[11] Kaplan, D.J. and Duncan, C.H. (1988) Nucleic Acids Res. 16, 10375.
[12] Isackson, P.J., Fishback, J.L., Bidney, D.L. and Reeck, G.R. (1979) J. Biol. Chem. 254, 5569–5572.
[13] Bianchi, M.E., Beltrame, M. and Paonessa, G. (1989) Science 243, 1056–1059.