Acta Chemica Scandinavica B 37 (1983) 731–753

Short Communications

Sequence Homology between Tissue Polypeptide Antigen (TPA) and Intermediate Filament (IF) proteins

BJÖRN LÜNING and ULRIKA NILSSON

Department of Organic Chemistry, Arrhenius-laboratoriet, University of Stockholm, S-106 91 Stockholm, Sweden

Partial sequences of Tissue Polypeptide Antigen (TPA) a protein isolated from human carcinomas or human placenta [1] have been published.[2] Comparisons with other proteins did not show any homology or analogy.

Recently the total structure of desmin laid the basis for a comparison covering epidermal keratins, hard keratins, vimentin and neurofilament proteins, showing these to form a class of filamentous proteins (IF proteins) with similar architecture.[3–5] Homology reaches nearly 80 % in some conservative regions and the secondary structure is based on three helices, Helix Ia, Helix Ib and Helix II, separated by spacers, a 6K headpiece and a tail of varying length and structure. The partial sequences of TPA when aligned with those of the IF proteins show an extensive homology (Figs. 1–3). Between TPA fragment BrCN:B and human epidermal 50K keratin there is 71 % homology (Fig. 1). In this region, which covers most of helix Ia, an epitope is located since the synthetic peptides 64 and 118

```
B    —— MAVLNDRLAQYLDEVRALEAANG-LEVL ————————
64        TDLDDRLAKYLDKVRALEAADGELGV
118       ALLNDELAEYLALVRALEAADGKLGV
7    ———————— RFAAFIDKVR ————————————————————
8      59 MQFLNDRLASYLEKVRQLERENAELESRILER.....
E      56 MQNLNDRLASYLDKVRALEEANADLEVKIRDW.....
D     103 LQELNDRFANYIEKVRFLEQQNALMVAEVNRL.....
GFA ———— MLNEEFARYIERVVFLEEQKRARAALLDE————
V      89 LQELNDRFADYIDKVRFLEQQNKILLAELEQL.....
```

Fig. 1. The amino acid sequence of TPA BrCN:B fragment and synthetic peptides 64 and 118 aligned with IF proteins. TPA BrCN:B fragment (B), Synthetic peptides 64 (64) and 118 (118), sheep wool α-keratin 7c (7), Sheep wool α-keratin 8c-1 (8), human epidermal 50K keratin (E), chicken gizzard desmin (D), porcine vimentin (V), bovine glial fibrillary acidic protein (GFA). Sequences of IF proteins are from Ref. 5, where further references are found. Solid lines denote unknown sequences.

```
E1   ———— MD-IIAE-VKAQYED-A-RM ————————————————
F    ——————————————————————————MLEE(F V)-L(F W)————
7    248 DLNMDCIVAEEIRAQYDDIASRSRAEAESWYRSK....
8    219 DLN---RVLNETRAQYEALVETNRRDVEEWYIRQ....
E    216 DLS---RILNEMRDQYEKMAEKNRKDAEEWFFIK....
D    260 DLT---AALRDVRQQYESVAAKNLQEAEEWYKSK....
V    246 DLT---AALRDVRQQYESVAAKNLQEAEEWYKSK....
```

Fig. 2. The amino acid sequences of TPA BrCN fragments E_1 and F aligned with IF proteins. TPA BrCN:E_1 (E_1), TPA BrCN:F (F), sheep wool α-keratin 7c (7), sheep wool α-keratin 8c-1 (8), human epidermal 50K keratin (E), chicken gizzard desmin (D), porcine vimentin (V). Sequences for IF proteins are from Ref. 5, where further references are found.

0302-4369/83 $2.50

Fig. 3. The amino acid sequences of TPA BrCN:C and synthetic peptides 107, 255 and 269 aligned with IF proteins. TPA BrCN:C (C), peptides 107 (107), 255 (255), and 269 (269), sheep wool α-keratin 7c (7), sheep wool α-keratin 8c-1 (8), human epidermal 50K keratin (E), chicken gizzard desmin (D), porcine vimentin (V), porcine neurofilament 68 K protein (NF). Sequences of IF proteins are from Ref. 5, where further references are found. Double headed arrow denotes the approximate location of the epitope in TPA BrCN:C.

are recognized by the antibody to TPA. Arginine is essential for the binding[6] and since the only arginine located at the same place in the protein and the synthetic peptides is the one corresponding to E 71 the epitope should be centered there. The high homology in this area will also explain why TPA antibody stains various epithelia in addition to carcinomas. The homology in this area with non-epithelial proteins is lower (30–40 %).

Assuming further homology between TPA and the 50K epidermal keratin, the nearest methionine would give an 11K fragment, which is very nearly the weight observed for the BrCN:B fragment.

The fragment BrCN:E_1 can be aligned with the IF proteins around the start of Helix II (Fig. 2). The TPA sequence shows homology with different proteins in the group totalling 66 %. During sequencing BrCN:E_1 no amino acids are recorded after 17 steps. Assuming a methionine at step 18, fragment BrCN:F, which has been proven to contain the single tryptophan residue of TPA[6] can be inserted after the BrCN:E_1 fragment, giving the tryptophan residue the "right" position (Fig. 2). Further along Helix II the TPA fragment BrCN:C can be aligned (Fig. 3). After sequencing the fragment for 56 steps no further amino acids could be detected. Aligned with the IF proteins four of them have methionine at the next step, indicating that the TPA fragment also should be terminated here, although it was originally believed to be somewhat larger.

The pattern of homology between this fragment and the other IF proteins is different from that in the previous two examples (Fig. 3). The TPA sequence borrows features from all the other IF proteins and an overall homology with the group is 46 %. However, the homology is low, only 9 % in the central third of the fragment and approaches 70 % in the two terminal thirds.

Synthetic peptides matching the natural sequences were tested by hemagglutination and RIA methods in an attempt to reveal the position of the epitope in BrCN:C. The peptides 255 and 269 are recognized by the TPA antibody whereas peptide 107 is not. The epitope should therefore be situated in the area of overlap between peptides 255 and 269, which is also the region of low homology with other IF proteins. The epitope located in TPA fragment BrCN:C could thus be expected to be more specific for the protein than the epitopes in the BrCN:B fragment. Arginine is essential for the ability to bind to the antibody,[6] and around arginine 31 in BrCN:C there is an area of low probability for α-helix formation. However, the coiled coil

heptade structure is retained.

The earlier ideas about a triple stranded architecture of the IF proteins has been doubted by Weber[3] who found evidence for a double stranded coiled coil.[3] The sedimentation analysis of TPA[1] confirms this observation as TPA forms a rod like dimer at high and low pH (2.1*S*, f/f_o=2.4). Around pH 7 soluble 4*S* aggregates are formed, which however, easily precipitate, properties intermediate between those observed for the desmin rod and the complete desmin molecule.[3]

Based on sequence homology between various parts of the TPA molecule and a number of IF proteins TPA should belong to that group and immunological cross reactions can be expected. However chemical differences in at least one epitope may explain immunological differences between TPA and other IF proteins.

Experimental conditions for the preparation of TPA and its antibody, as well as the testing of antibody-antigen binding have been presented[1,2,6]. Peptides were synthesized by Merrifield technique.[7]

Acknowledgement. Thanks are due to Prof. Klaus Weber for valuable discussions and for giving access to Refs. 4 and 5 prior to their publication.

Note added in proof. A recently published Type II human epidermal cytoskeletal keratin [Hanukoglu, I. and Fuchs, E. *Cell 33* (1983) 915] shows a high homology (60 %) with the TPA BrCN:C fragment although TPA is a Type I cytokeratin.

1. Lüning, B., Wiklund, B., Redelius, P. and Björklund, B. *Biochim. Biophys. Acta 624* (1980) 90.
2. Redelius, P., Lüning, B. and Björklund, B. *Acta Chem. Scand. B 34* (1980) 265.
3. Geisler, N. and Weber, K. *EMBO J. 1:12* (1982) 1649.
4. Geisler, N., Kaufmann, E., Fischer, S., Plessmann, U. and Weber, K. *EMBO J. 2:8* (1983). *In press.*
5. Weber, K., Shaw, G., Osborn, M., Debus, E. and Geisler, N. *Cold Spring Harbor Symp. Quant. Biol. 48* (1983). *In press.*
6. Wiklund, B., Lüning, B. and Björklund, B. *Acta Chem. Scand. B 35* (1981) 325.
7. Ragnarsson, U. and Lüning, B. *Unpublished results.*

Received August 18, 1983.