Differential proteolysis and evidence for a residue exchange in tissue plasminogen activator suggest possible association between two types of protein microheterogeneity

Hans Jörnvall, Gunnar Pohl, Nils Bergsdorf and Per Wallén

Department of Physiological Chemistry, Karolinska Institutet, S-104 01 Stockholm and *Umeå University, S-901 87 Umeå, Sweden*

Received 31 March 1983

The N-terminal part of native one-chain tissue plasminogen activator from melanoma cells is not homogeneous. The protein chain starts at two different positions, in all probability representing a processing difference in the N-terminus. Both 'long' L-chains and 3-residue shorter S-chains are present in the preparations. In addition, results compatible with a positional Ser/Gly microheterogeneity were obtained at a single position (position L-4 which is equal to S-1). The N-terminal tripeptide difference seems to be coupled to the possible microheterogeneity: L-chains contain Ser in this position, while S-chains appear to contain predominantly Gly.

N-terminal proteolysis *Protein processing* *Protein heterogeneity* *Amino acid sequence analysis*

1. INTRODUCTION

Tissue plasminogen activator has been purified from culture medium of a melanoma cell line, Bowes [1,2]. It is a one-chain glycosylated polypeptide with M_r ~72000 [1–3], for which a 527-residue amino acid sequence has been deduced from analysis of the corresponding cDNA [4]. A shorter cDNA segment has also been cloned [5]. On treatment with plasmin, the one-chain activator is cleaved approximately in the middle [2,3] to produce 2 disulfide-connected chains (A and B, from the N- and C-terminal ends, respectively). The B-chain is homologous to the catalytic chains of plasmin, thrombin and other serine proteases [2,4,5]. Properties of the different activator forms are summarized in fig.1, which also explains the nomenclature used here.

Apart from the one-chain/two-chain variation, 3 other types of possible variants have been reported [2]. First, direct sequence analysis revealed that the molecules consist of polypeptides starting at different positions [2]. Thus, the material from melanoma cell cultures can exhibit two N-terminal sequences, differing by the lack of the tripeptide H_2N-Gly–Ala–Arg– at the start of one of the structures. Consequently, the shorter chain (S) starts with position 4 of the longer variant (L). Preparations apparently exhibiting only S-chains have also been reported [4]. Plasmin (apart from giving the two-chain cleavage) cleaves off the extra N-terminal tripeptide in the L-chain [2], thus removing the N-terminal S/L heterogeneity in the A-chain. Therefore, the A-chain is always S-type.

A second variation is reflected by molar mass differences, possibly due to internal carbohydrate differences in the A-part, constituting type I and type II activator molecules [4,6]. Finally, some evidence for positional microheterogeneities has been obtained. A Ser/Gly exchange at position S-1 appeared possible [2], but results on analyses of S-, L- and A-chains were partly conflicting. The evidence for the residue exchange is now repeatedly obtained. In addition, the different results on the conflicting identity of the residues can be explained and seem to correlate with the S/L dif-

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

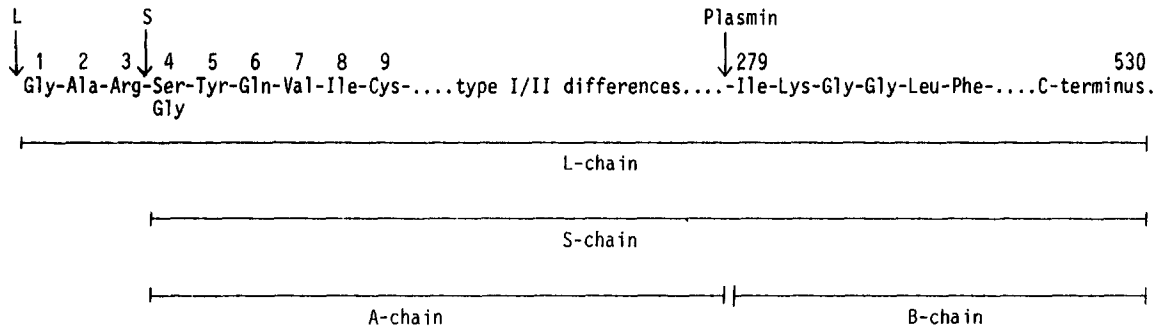

Fig.1. Schematic representation of the different forms of tissue plasminogen activator. Four different types of variation seem to occur in activator preparations: (a) one-chain forms (L + S-chains) versus two-chain forms (disulfide-linked A and B chains, as obtained by cleavage during activation by plasmin); (b) internal differences giving types I and II (with different M_r); (c) processing (at arrows marked L and S) differences (with non-stoichiometric cleavages at S) giving L- and S-chains (plasmin can also complete cleavage at S apart from cleaving between A- and B-chains); and (d) suggested positional differences at L-4. (a,b) relationships are from [2–4]; (c,d) results from this study. Positional numbers refer to the L-chain (with B-chain numbers taken from [4] after adjustment to the L-chain relationships now reported).

ference, thus suggesting an apparent association between two types of microheterogeneity in the activator protein.

2. MATERIALS AND METHODS

Tissue plasminogen activator was prepared and treated with plasmin as in [2]. For purification of the N-terminal CNBr-fragment, native preparations were first treated with plasmin, reduced and carboxymethylated [2]. The carboxymethylated A- and B-chains (40 nmol) were then separated on an arginine–Sepharose column (2.5 × 3 cm) equilibrated with 0.5 M ammonium bicarbonate and 0.01% Tween 80. The B-chain eluted unretarded, while the bound A-chain was eluted with guanidinium hydrochloride (0.6 M). After desalting, the carboxymethylated A-chain (25 nmol) was treated for 24 h with CNBr (125 mg) in 70% formic acid (1 ml). Three fragments were separated on a 1.6 × 150 cm TSK 55F column (Merck), equilibrated with 30% acetic acid, the smallest being the N-terminal CNBr-fragment corresponding to positions 4–16 of the native L-chain (positions 1–13 of the S-chain).

Total compositions were determined by amino acid analysis in a Beckman 121M analyzer after acid hydrolysis [2]. Amino acid sequences were determined by liquid-phase sequencer degradations (Beckman 890D), using a 0.1 M Quadrol peptide program after application in 1% sodium dodecyl sulfate (for one-chain preparations) or into precycled polybrene [7]. Manual degradations were carried out with the dimethylaminoazobenzene isothiocyanate method [8], utilizing by-products in the identifications [9].

3. RESULTS

Both native (one-chain) and plasmin-treated (two-chain) activator preparations revealed two sequences on direct sequence analysis. The structural explanation is given in fig.1, and the supporting results of direct sequence analysis are summarized in table 1. For the native one-chain activator, two residues, corresponding to the L- and S-chains, were found in largely similar amounts in all but the first position (table 1). The recovery of L- and S-chain residues varied somewhat between preparations, suggesting L/S ratios of between 1:3 and 1:1 in 5 different preparations studied, including separated type I and type II one-chain activators.

The plasmin-treated preparations also gave 2 sequences, corresponding to the A- and B-chains, but the A-chain sequence gave only single results (corresponding to the S-type), except for cycle 1, where two residues (Ser and Gly) were obtained (table 1). The direct sequence analyses therefore appear to suggest that native single-chain preparations predominantly consist of protein chains of the type L_{Ser-} and S_{Gly-1}, recovered in similar amounts. Manual sequence analysis of the N-

Table 1

Results of sequence analyses

Cycle	One-chain preparation		Two-chain preparation		N-terminal CNBr-fragment
	L-chain	S-chain	A-chain	B-chain	
1	Gly (3.6)	Gly (3.6)/Ser (trace)	Ser (6)/Gly (16)	Ile (25)	Ser Gly
2	Ala (4.4)	Tyr (5.2)	Tyr (19)	Lys (18)	Tyr
3	Arg (3.0)	Gln (1.9)	Gln (14)	Gly (15)	Gln
4	Ser (2.4)	Val (5.5)	Val (23)	Gly (13)	Val
5	Tyr (2.2)	Ile (4.2)	Ile (21)	Leu (23)	Ile
6	Gln (1.9)	Cys(Cm) (2.2)	Cys(Cm) (6)	Phe (18)	Cys(Cm)

One-chain (L + S) preparations (first column) and two-chain (A + B) preparations (second column) analyzed by liquid-phase sequencer degradations, and N-terminal CNBr-fragment analyzed by manual dimethylaminoazo benzene isothiocyanate degradations. Results from mixtures are listed ordered into the structure to which they belong, values show nmol recovered from the high-performance liquid chromatography identifications (Gly, Ser, Arg, Gln, Cys(Cm) normally recovered in lower yield than hydrophobic residues). Nomenclature is given in fig.1

terminal CNBr-fragment from the A-chain supported this conclusion by showing the presence of N-terminal Gly and Ser in equal amount.

Combined, the results reveal that of the 4 theoretically possible chains that may be derived from one processing difference (L/S chains) and one apparent positional microheterogeneity (Gly/Ser), only two chains seem to exist in larger quantity, $L_{Ser\text{-}4}$ and $S_{Gly\text{-}1}$ (cf. fig.1), suggesting an association between the two types of heterogeneity.

4. DISCUSSION

One heterogeneity in the N-terminal region of native melanoma cell one-chain tissue plasminogen activator is clearly shown (table 1). Hence, protein chains of 2 types (L and S) exist, differing by the presence or absence of 3 residues. Interestingly, another such difference has been noticed in a completely different activator [10]. Relative amounts of the present melanoma activator's L and S chains vary but both chains have been detected in 5 different preparations studied. The presence of an N-terminal heterogeneity in starting position is in agreement with another report [2]. However, other preparations of the same melanoma cell activator have apparently been found to start at only one position (present S-type of chain) [4]. This variability in starting position may be due to inherited differences between cell cultures or to the presence of proteolytic (e.g., plasmin-like) contaminants, cleaving the arginyl bond after residue 3 of the L-chain. In any event, cell cultures can yield different preparations, and a starting position previously deduced [4] need not be characteristic for tissue plasminogen activator.

The other heterogeneity discussed, concerns a suggested Gly/Ser exchange at position L-4 (= S-1). Sequence analysis of the N-terminal CNBr-fragment suggests both Gly and Ser in the position corresponding to S-1 (table 1), and direct sequence analyses of native activator preparations reveal essentially only Gly (in the first cycle corresponding to L-1 and S-1) but essentially only Ser in the fourth cycle corresponding to L-4 (plus Val from S-4, cf. table 1). Similarly, plasmin-treated preparations appear to give both Ser and Gly from position A-1 (plus Ile from B-1).

The positional heterogeneity is consistent with results from one study [2] but was not reported from another [4]. Therefore, it is possible, exactly as with the processing, that different cultures may produce different structures. Of course, the discovery of Gly in S-1 and A-1 but Ser in L-4 can reflect a contamination with free Gly. Data are also difficult to interpret since all preparations are anyway heterogeneous from the L/S heterogeneity (also affecting Gly in the first cycle). However, a Gly-contamination seems insufficient as explanation to the results, since not only is Gly detected in cycle 1 of the intact activator (table 1, column 1)

but Ser is also consistently recovered in a lower amount in this cycle than in cycle 4 of the native preparation (table 1, column 1) or in cycle 1 of the plasmin-cleaved preparation (table 1, column 2). Therefore, if N-terminal Ser is not destroyed and a Gly contamination simultaneously acquired in the same preparations, the results are difficult to explain in the absence of heterogeneity in position L4/S1. It is true that N-terminal Ser may be preferentially removed by selective acetyl-group blocking during nascent processing [11], but a coincidence between such a blocking and a Gly-contamination also seems unlikely. In conclusion therefore, a positional Gly/Ser microheterogeneity is the probable explanation to the results obtained. The origin of this heterogeneity is unknown. However, the presence of a second genetic locus seems unlikely [4]. Possibly therefore, the heterogeneity may be derived from normal allelic variations or from mutational events after separation of the Bowes cultures studied.

Finally, the suggested positional Gly/Ser difference appears to correlate with the 3-residue L/S shift. Thus, of the 4 structural combinations possible from 2 different heterogeneities, only 2 are detectable, suggesting an uncommon association between 2 types of protein heterogeneity. This pattern explains the previously conflicting data obtained on analyses of S-, L- and A-chains [2].

Regarding the origin of the association, it may be noticed that the activator is synthesized in a larger pro-form [4] which is released to the culture medium after proteolytic cleavage(s) in the N-terminus. Two explanations to the association of heterogeneities appear possible. Either, pro-forms may differ in additional positions that affect the signal, and hence position, for cleavage. Alternatively, the protease responsible for the cleavage(s) may prefer Arg–Gly bonds before Arg–Ser bonds. Interestingly, both these specificities are different from those typical for nascent signal peptide cleavages [12]. Instead, the Arg–Gly cleavage (which is from a structure Arg–Arg–Gly [4]) resembles the specificity of hormonal and other pro-form cleavages [4] which typically occur after dibasic structures [13]. That specificity even appears common enough to be reflected as a partly restricted sequence in proteins not destined for this cleavage [14]. In any event, the two-chain processing by plasmin is obviously different from the N-terminal L/S processing, since the L/S difference created in the one-chain processing is destroyed by plasmin. Independent of the processing enzymes, the present results establish complicated patterns in the activator N-terminal region, and suggest associations between different events.

ACKNOWLEDGEMENTS

This work was supported by grants from the Swedish Medical Research Council (projects 13X-3532 and 13X-3906), the Magn. Bergvall's Foundation, the Swedish Cancer Society, the Knut and Alice Wallenberg's Foundation, and KabiGen. Skilful technical assistance by Margareta Källström, Carina Palmberg, Rut Nyström and Maria Swärd is gratefully acknowledged.

REFERENCES

[1] Rijken, D.C. and Collen, D. (1981) J. Biol. Chem. 256, 7035–7041.
[2] Wallén, P., Pohl, G., Bergsdorf, N., Rånby, M., Ny, T. and Jörnvall, H. (1983) Eur. J. Biochem., in press.
[3] Wallén, P., Rånby, M., Bergsdorf, N. and Kok, P. (1981) Prog. Chem. Fibrinolysis Thrombolysis 5, 16–23.
[4] Pennica, D., Holmes, W.E., Kohr, W.J., Harkins, R.N., Vehar, G.A., Ward, C.A., Bennett, W.F., Yelverton, E., Seeburg, P.H., Heyneker, H.L., Goeddel, D.V. and Collen, D. (1983) Nature 301, 214–221.
[5] Edlund, T., Ny, T., Rånby, M., Hedén, L.-O., Palm, G., Holmgren, E. and Josephson, S. (1983) Proc. Natl. Acad. Sci. USA 80, 349–352.
[6] Rånby, M., Bergsdorf, N., Pohl, G. and Wallén, P. (1982) FEBS Lett. 146, 289–292.
[7] Jörnvall, H. and Philipson, L. (1980) Eur. J. Biochem. 104, 237–247.
[8] Chang, J.Y., Brauer, D. and Wittman-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[9] Von Bahr-Lindström, H., Hempel, J. and Jörnvall, H. (1982) J. Prot. Chem. 1, 257–261.
[10] Günzler, W.A., Steffens, G.J., Ötting, F., Kim, S.-M.A., Frankus, E. and Flohé, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1155–1165.
[11] Jörnvall, H. (1975) J. Theor. Biol. 55, 1–12.
[12] Austen, B.M. (1979) FEBS Lett. 103, 308–313.
[13] Pradayrol, L., Jörnvall, H., Mutt, V. and Ribet, A. (1980) FEBS Lett. 109, 55–58.
[14] Jörnvall, H. and Persson, B. (1983) Biosci. Rep. 3, 225–232.