# THE SUBUNITS OF DNA-DEPENDENT RNA POLYMERASE FROM *E. COLI*: I. AMINO ACID ANALYSIS AND PRIMARY STRUCTURE OF THE N-TERMINAL REGIONS

Hirota FUJIKI and Gabriela ZUREK
*Max-Planck-Institut für Biochemie, 8033 Martinsried bei München, West Germany*

Received 20 May 1975

## 1. Introduction

Recent progress in the understanding of the function of the DNA-dependent RNA polymerase subunits in transcription makes the elucidation of the structural analysis of the subunits a necessity. We here report the quantitative determinations of the amino acid compositions and the amino acid sequences of the N-terminal region of all four subunits of the polymerase as a basis for the primary structure. So far, the amino acid composition of $\sigma$ factor and of the separated $\beta$ and $\beta'$ subunit have not been determined.

## 2. Materials and methods

### 2.1. *RNA polymerase purification and subunit separation*

The purification of *E. coli* RNA polymerase from *E. coli* K 12 and from AJ7, rif-rJ7 was done as previously described [1]. The holoenzyme was further chromatographed on phosphocellulose to isolate pure core enzyme and $\sigma$ factor [2]. The $\alpha$, $\beta$ and $\beta'$ subunits of wild type *E. coli* were further separated from pure core enzyme by electrophoresis of preparative cellulose acetate blocks in 6 M urea, 0.6 M boric acid buffer containing 0.01 M EDTA, 0.02 M mercaptoethanol, pH 8.9. The $\sigma$ factor of the AJ7-enzyme was separated by electrophoresis on cellulose acetate blocks in 6 M urea, 0.2 M ammonium bicarbonate buffer, containing 0.01 M magnesium acetate, 0.001 M EDTA, 0.02 M mercaptoethanol, pH 9.0 [3].

The purity of the individual subunits was checked by sodium dodecylsulfate (SDS) polyacrylamide gel electrophoresis [4].

### 2.2. *Amino acid analysis*

A 50% yield of the $\alpha$, $\beta$ and $\beta'$ subunits was obtained. The amino acid compositions of these three subunits were determined with a conventional 'Unichrom' (Beckman, Munich) amino acid analyzer.

The yield of $\sigma$ factor obtained by the electrophoretic procedure is low in comparison with that of the other subunits. For the amino acid analysis of $\sigma$ factor the 'Unichrom' was modified by the installation of a micro volume dual channel colorimeter from LKB-Biocal. This allowed an amino acid analysis in the 1–10 nmol range. The modification of the amino acid analyzer will be reported in another paper. [5].

### 2.3. *Conditions for HCl-hydrolysis*

The sample of protein was hydrolysed in constant boiling 5.7 N hydrochloric acid at 110°C for 20 hr. The acid was then removed under vacuum at 50°C.

For the determinations of cysteine, half cystine and methionine, samples were first oxidized with performic acid [6]. For the determination of tryptophan , the protein was hydrolysed with 5.7 N HCl containing 6% thioglycolic acid, or with 3 N *p*-toluenesulfonic acid containing 0.2% 3-(2-amino ethyl) indole [7,8].

### 2.4. *Micro Edman degradation procedure for proteins*

A 10 nmol protein sample was heated in a solution of SDS at 80°C for 5 min. After evaporating to dryness the manual Edman degradation was performed according to a modified micro procedure [9].

Phenylthiohydantoin (PTH) amino acids were identified on polyamide layers (5 × 5 cm) by two-dimensional chromatography using butyl-PBD fluorescence indicator which was a gift of Pierce Chemicals [10].

 *North-Holland Publishing Company – Amsterdam*

## 3. Results and discussion

### 3.1. *Amino acid composition of the subunits*

The amino acid composition of the monomeric core enzyme has been determined by two groups [11,12]. Their reported molar contents of the amino acids with the exception of the methionine residue are very similar.

R. Burgess has reported the amino acid composition of the $\alpha$ subunit and of a mixture of the $\beta$ and $\beta'$ subunits [13]. By preparative electrophoresis on a cellulose acetate block in 6 M urea, holoenzyme has now been fractionated into the $\alpha$, $\beta$ and $\beta'$ subunits and $\sigma$ factor. This technique was the first important step for the study of the proteinchemistry of the subunits of RNA polymerase.

The electrophoretic migration of all subunits on cellulose acetate at pH 8.9 is from cathode to anode in the order $\beta'$, $\beta$, $\alpha$ and $\sigma$. Therefore, $\sigma$ factor is the most acidic and $\beta'$ subunits the most basic protein. Table 1 shows the amino acid residues in molar percentage. The $\beta'$ subunits contain the highest molar percentage of lysine, histidine and arginine taken altogether. The $\sigma$ factor possesses a very high percentage of aspartic and glutamic acid residues. However, the $\sigma$ factor can be suspected to contain many aspartic acid and glutamic acid residues rather than asparagine and glutamine because it is the most acidic protein. The molecular weight of the $\alpha$, $\beta$ and $\beta'$ subunits and $\sigma$ factor determined by SDS-polyacrylamide gel electrophoresis are 40 000, 155 000, 160 000 and 92 000 respectively [14]. On the basis of these molecular weights the nearest integral number of amino acid residues were calculated as shown in table 2.

Table 1
Amino acid composition of each subunit (mole %)

| Amino acid | $\alpha$ | $\beta$ | $\beta'$ | $\sigma$ |
|---|---|---|---|---|
| Lysine | 4.79 | 5.60 | 6.12 | 5.58 |
| Histidine | 2.10 | 1.57 | 1.51 | 1.99 |
| Arginine | 6.54 | 6.83 | 7.34 | 6.51 |
| Cysteic acid | 1.11 | 0.70 | 1.00 | 0.93 |
| Aspartic acid | 9.18 | 10.25 | 9.43 | 13.15 |
| Methionine sulfone | 1.34 | 2.27 | 2.30 | 4.90 |
| Threonine | 5.61 | 4.90 | 5.69 | 6.27 |
| Serine | 5.90 | 5.95 | 5.40 | 4.28 |
| Glutamic acid | 13.91 | 13.75 | 11.67 | 18.36 |
| Proline | 4.67 | 4.38 | 4.25 | 3.54 |
| Glycine | 7.30 | 7.88 | 8.35 | 4.47 |
| Alanine | 7.07 | 6.48 | 8.71 | 5.27 |
| Valine | 9.29 | 7.88 | 6.98 | 5.21 |
| Isoleucine | 6.25 | 5.69 | 5.61 | 6.27 |
| Leucine | 10.81 | 9.72 | 9.94 | 8.00 |
| Tyrosine | 1.75 | 2.80 | 2.52 | 2.11 |
| Phenylalanine | 1.57 | 2.89 | 2.59 | 2.48 |
| Tryptophan | 0.70 | 0.35 | 0.50 | 0.68 |

Table 2
Nearest integral number of amino acid residues of each subunit

| Amino acid | $\alpha$ | $\beta$ | $\beta'$ | $\sigma$ |
|---|---|---|---|---|
| Lysine | 17 | 79 | 92 | 47 |
| Histidine | 8 | 22 | 23 | 17 |
| Arginine | 24 | 96 | 110 | 54 |
| Cysteic acid | 4 | 10 | 15 | 8 |
| Aspartic acid | 33 | 144 | 141 | 110 |
| Methionine sulfone | 5 | 32 | 35 | 41 |
| Threonine | 20 | 69 | 85 | 52 |
| Serine | 21 | 84 | 81 | 36 |
| Glutamic acid | 51 | 194 | 175 | 153 |
| Proline | 17 | 62 | 64 | 30 |
| Glycine | 27 | 111 | 125 | 37 |
| Alanine | 26 | 91 | 131 | 44 |
| Valine | 34 | 111 | 105 | 44 |
| Isoleucine | 23 | 80 | 84 | 52 |
| Leucine | 39 | 137 | 149 | 67 |
| Tyrosine | 6 | 39 | 38 | 18 |
| Phenylalanine | 6 | 41 | 39 | 21 |
| Tryptophan | 3 | 5 | 8 | 6 |
| Total | 364 | 1407 | 1500 | 837 |

### 3.2. *The N-terminal sequences of the subunits*

The N-terminal residues of $\alpha$, $\beta$ and $\beta'$ subunits and $\sigma$ factor are all methionine as shown by the method of dansylation [15] and by Edman degradation [16].

Table 3 shows the N-terminal sequences of all four subunits. Through the microscale procedure, PTH amino acid could be identified from a 10 nmol sample of protein. This method was especially suitable for the sequence of $\sigma$ factor. The fourth step of the $\beta'$ subunit was identified as leucine by back hydrolysis of the PTH amino acid with hydriodic acid. The hydrolysate was also dansylated because PTH leucine and PTH isoleucine could not be separated on polyamide

Table 3
N-terminal sequences of each subunit

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\alpha$ | Met – | GIN – | Gly – | AsN – |
| $\beta$ | Met – | Val – | Tyr – | Arg – |
| $\beta'$ | Met – | Lys – | Asp – | Leu – |
| $\sigma$ | Met – | Glu – | GIN – | AsN – |

layers by two-dimensional chromatography. It has been suggested that the $\beta$ and $\beta'$ subunits were gene-multiplicated products of the $\alpha$ subunit. Although the molecular weights of the subunits are very different, the values of the molar percentages of the amino acids seem to be relatively similar. This assumption will remain until an amino acid analysis of all the subunits has been performed. At first, we have compared the N-terminal sequences of each subunit over several steps. The N-terminal sequence of the $\alpha$ subunit is being continued further and will be reported on in another paper.

## Acknowledgements

The authors wish to thank Professors W. Zillig and G. Braunitzer, as well as the Sonderforschungsbereich 51 for generous support of this work.

## References

[1] Zillig, W., Zechel, K. and Halbwachs, H.-J. (1970) Hoppe-Seyler's Z. Physiol. Chem. 351, 221–224.
[2] Burgess, R. R., Travers, A. A., Dunn, J. J. and Bautz, E. K. F. (1969) Nature 221, 43–46
[3] Heil, A. and Zillig, W. (1970) FEBS Lett. 11, 165–168.
[4] Shapiro, A. L., Vinnela, E. and Maizel, J. V. Jr. (1967) Biochem. Biophys. Res. Comm. 28, 815–820.
[5] Fujiki, H. and Riess, W., manuscript in preparation.
[6] Hirs, C. H. W. (1956) J. Biol. Chem. 219, 611–621.
[7] Matsubara, H. and Sasaki, R. M. (1969) Biochem. Biophys. Res. Common. 35, 175–181.
[8] Liu, T. Y. and Chang, Y. H. (1971) J. Biol. Chem. 246, 2842–2848.
[9] Peterson, J. D., Nehrlich, S., Oyer, P. E. and Steiner, D. F. (1972) J. Biol. Chem. 247, 4866–4871.
[10] Summers, M. R., Smythers, G. W. and Oroszlan, S. (1973) Anal. Biochem. 53, 624–628.
[11] Priess, H. and Zillig, W. (1967) Biochim. Biophys. Acta 140, 540–542.
[12] Maitra, V. and Hurwitz, J. (1967) J. Biol. Chem. 242, 4897–4907.
[13] Burgess, R. R. (1967) J. Biol. Chem. 244, 6168–6176.
[14] Zillig, W., personal communication.
[15] Hartley, B. B. (1970) Biochem. J. 119, 805–822.
[16] Edman, P. (1956) Acta Chem. Scand. 10, 761–768.
[17] Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. and Ballantyne, D. L. (1971) Biochemistry 10, 4912–4921.