Cell, Vol. 14, 179–190, May 1978, Copyright © 1978 by MIT

Patterns of Protein Synthesis in E. coli: a Catalog of the Amount of 140 Individual Proteins at Different Growth Rates

Steen Pedersen,* Philip L. Bloch, Solvejg Reeh* and Frederick C. Neidhardt
Department of Microbiology
The University of Michigan
Ann Arbor, Michigan 48109

Summary

The amount of 140 individual proteins of E. coli B/r was measured during balanced growth in five different media. The abundance of each protein was determined from its absolute amount in ^{14}C-glucose-minimal medium and a measurement of its relative amount at each growth rate using a double labeling technique. Separation of the proteins was carried out by two-dimensional gel electrophoresis.

This catalog of proteins, combined with 50 additional ribosomal proteins already studied, comprises about 5% of the coding capacity of the genome, but accounts for two thirds of the cell's protein mass.

The behavior of most of these proteins could be described by a relatively small number of patterns. 102 of the 140 proteins exhibited nearly linear variations with growth rate. The remaining 38 proteins exhibited levels which seemed to depend more on the chemical nature of the medium than on growth rate.

Proteins, including the ribosomal proteins, that increase in amount with increasing growth rate account for 20% of total cell protein by weight during growth on acetate, 32% on glucose-minimal medium and 55% on glucose-rich medium. Proteins with invariant levels in the various media comprise about 4% of the cell's total protein.

Introduction

Studies of the regulation of protein synthesis have mostly been focused on specific enzymes where the enzymatic activity has been used as an assay for the amount of protein. Other studies have been possible only in the few cases where some physical property of the protein in question is distinctive. In this way, the regulation of the very large subunits of the E. coli RNA polymerase (Matzura, Hansen and Zeuthen, 1973) and of the small basic ribosomal proteins has been studied, in the latter case by the electrophoresis technique developed by Kaltschmidt and Wittmann (1970).

For lack of a suitable method, attention could not be given to the overall pattern of protein regulation in the cell. Earlier attempts to survey the cell's protein profile (Raunio and Savimo, 1965; Moses and Wild, 1969) lacked sufficient resolving power to reach the level of individual proteins. The first real opportunity to attempt a global depiction of the cell's pattern of protein synthesis has been provided by the O'Farrell (1975) technique for resolving complex protein mixtures on two-dimensional polyacrylamide gels.

This report presents the results of an ongoing analysis of the proteins of E. coli B/r using the O'Farrell technique. We have measured the amounts of a large number of proteins, some identified, but most unknown, during steady state growth in five defined media. Our goal has been to define groups of proteins that might share regulatory responses, if such groups exist, and also to learn how the activity of one protein might affect the biosynthesis of other proteins. Such knowledge is essential for discriminating among different models of regulation, and for designing proper in vitro experiments that can elucidate the detailed molecular mechanism of the regulation of each protein. Eventually one wants to understand how a given growth medium leads to the singular growth rate of E. coli cells in that medium.

Results

The proteins chosen for study were selected as recognizable spots in the two-dimensional array of total cell protein on O'Farrell gels prepared from cells grown in the five different media. This selection implies, in most cases, that the protein is abundant (>0.05% of total protein) at least at one growth rate. An example of the spot pattern can be found in Figure 1 in the Appendix describing the system of protein nomenclature used.

We have described the pattern of regulation of the majority of the cell's protein mass under different steady state growth conditions. This was done by determining the absolute amount of each of 140 proteins in one medium, followed by a determination of the relative amount of the protein at each growth rate. To avoid error caused by a different amino acid composition of the proteins, we measured the absolute amount of each protein in cells grown in minimal medium with uniformly labeled ^{14}C-glucose. The protein spots on these gels were cut out in their entirety, and their radioactivity was measured and normalized to the radioactivity in EF-G. The fraction of total protein represented by each protein (α') was calculated using the assumptions that one EF-G molecule is present per ribosome (Gordon, 1970), that the fraction of total protein which is ribosomal proteins is 0.137 [determined in this strain in glucose medium by Dennis

* Present address: Department of Molecular Biology, University of Aarhus, DK-8000 Aarhus C, Denmark.

and Bremer (1974)] and that the recovery of each protein is similar to that of EF-G. Recovery of EF-G was previously determined to be about 0.94 (Neidhardt et al., 1977). The α' measured is equivalent in these steady state cultures to the α defined by Schleiff (1967) if one assumes that no significant turnover or selective loss of these proteins is occuring. The α' for each protein in glucose medium is given in Table 1.

The assumption that recovery of each spot is as high as that of EF-G might not be true for all proteins, and therefore these estimates should be taken as minimum values. That the recovery of proteins in general is satisfactory is supported by the observation that most of the proteins in the lysis mixture, exclusive of ribosomal proteins, actually enter the gel and appear as resolved spots (data not shown), and that the amount of protein in each spot relative to that of other spots remains constant from one gel to another.

The measurement under different growth conditions of the relative amounts of proteins was carried out in two sets of experiments. In one set, portions of a reference culture (glucose medium) labeled with ^{14}C-leucine were mixed prior to sonic treatment with each of the four cultures grown at different growth rates in the other media containing ^{3}H-leucine. Gels were prepared from each cell mixture, and the ^{3}H/^{14}C ratio of the total proteins and of each spot was measured. Labeling with the same amino acid has the advantage that each spot in the mixture of ^{3}H- and ^{14}C-labeled cells prepared on the same glucose medium should theoretically have a ^{3}H/^{14}C ratio identical to that of total protein. For the majority of proteins, this criterion was met within experimental error. Exceptions were found in spots B21.8, E106, F24.5, G31.6, G42.5, G49.2 and G74.0, indicating either that the two glucose cultures were not truly equivalent, or that problems existed with the processing and/or counting of these proteins. In the other set of experiments, the reference culture (glucose medium) was labeled with a mixture of ^{3}H-leucine plus ^{3}H-isoleucine and mixed prior to sonic treatment with cells growing at each of the various growth rates labeled with $^{35}SO_4$. The ^{3}H/^{35}S ratio in total protein and in each spot was determined. In this set of experiments, it was possible on the autoradiograms of the gels to detect spots which are prominent only in one of the nonreference growth conditions, but virtually absent in the glucose reference culture. The fastest growth rate at which labeling with $^{35}SO_4$ can be easily carried out is in the fully supplemented medium lacking methionine and cysteine, giving a maximal growth rate of 1.5 hr^{-1} instead of 2.0 hr^{-1}.

The purity of a protein spot was tested by measuring the ratio of (met + cys) to (leu + ile) for each growth condition by labeling cells in separate cultures of the same medium with $^{35}SO_4$ and ^{3}H (leucine and isoleucine) and then mixing the cells. This ratio, characteristic for each protein, ranged from 0.3–2 times the ratio in total protein. If a spot is composed of two or more proteins, this ratio will vary with growth rate if the proteins are regulated differently and have different isotope ratios. About ten relatively minor spots showed such variations, and the data for them are not presented.

Excellent agreement was found in the relative amount of each protein measured by the two labeling methods, except for eight spots that were probably misidentified and therefore have not been included. The reproducibility of double determinations was usually well within 10%. The level of each protein relative to the level in glucose medium is given in Table 1 for 140 proteins. Where the identity of the protein has been established, this is also shown in the table. The detailed criteria of identity are presented in the Appendix. The results obtained for A13.0 (L7), B13.0 (L12) and C14.8 (S6) are in excellent agreement with earlier results (Deusser, 1972; Dennis, 1974).

The identifying numbers for each protein listed in Table 1 provide the apparent molecular weight ($\times 10^{-3}$) of each protein as measured by its migration in the SDS-acrylamide gel (see Appendix). The number of molecules of each protein per genome in cells growing in glucose medium is shown in Table 1 together with the method of arriving at these estimations. The data of Table 1 were also used to construct graphs of the levels of each individual protein as a function of the growth rate (not shown). To organize this mass of data, we first assembled the proteins in groups showing similar regulation. Three groups exhibit a monotonic variation with growth rate. Group Ia (33 proteins) decreases in level, group Ib (16 proteins) is constant and group Ic (54 proteins) increases in level with increasing growth rate. The remaining proteins are classified as showing either a maximum (group IIa, 32 proteins) or a minimum (group IIb, 5 proteins) level at one of the intermediate growth rates. These groups of proteins, given in Table 1, can of course be subdivided further, as is discussed below. The weight fractions of all proteins at each growth rate were calculated, and the summed α' for each group is shown in Figure 1 as a function of the growth rate.

Discussion

The total α' for all proteins measured in this study ranged from 0.53 in acetate to 0.46 in rich medium. Adding the α' of unrecovered ribosomal proteins, which have isoelectric points outside the pH range

Table 1. Amount of Various Proteins in E. coli NC3 in Different Media

Protein Number[a]	Protein Identification[b]	Regulation Group[c]	Weight Fraction of Total Protein in Glucose Medium ($\alpha' \times 10^3$)	Number of Molecules per Genome in Glucose Medium[d]	Level in Each Medium Relative to the Level in Glucose Medium[e]			
					Acetate (k = 0.38)	Glycerol (k = 0.77)	Rich-Met (k = 1.50)	Rich (k = 1.98)
A13.0	L7	Ic	2.93	9,920	0.87	0.99	1.11	–
B13.0	L12	Ic	4.57	15,470	0.24	0.75	1.66	–
B14.6		Ic	0.68	2,050	0.77	1.04	1.05	–
B21.8		Ic	0.90	1,820	0.61	0.86	–	1.60
B26.3		IIa	4.83	8,080	0.87	1.13	–	0.35
B40.7	α	Ic	3.66	3,960	0.76	0.96	1.02	1.44
B46.7		Ia	5.64	5,310	1.15	1.19	0.72	0.94
B50.3		Ic	6.15	5,380	0.57	0.90	1.14	1.57
B56.5	A	Ic	13.5	10,510	0.94	1.00	1.18	1.22
B58.3		IIa	3.11	2,350	0.33	1.30	0.46	–
B61.0		Ic	1.25	900	0.59	0.67	1.20	–
B65.0	S1	Ic	15.1	10,220	0.58	0.84	1.33	1.74
B66.0		Ic	9.33	6,220	0.67	0.89	–	1.51
B82.5		Ia	0.28	150	2.88	1.95	1.14	–
B83.0		Ia	0.95	500	2.45	2.21	1.00	–
B88.0		Ia	1.45	720	1.11	1.00	0.91	–
B89.0		Ib	0.95	470	1.05	0.87	1.05	–
C14.8	S6	Ic	0.95	2,820	0.58	0.84	1.65	–
C15.3		Ic	1.24	3,570	0.57	0.94	1.28	–
C15.4		Ib	1.89	5,400	0.91	1.00	1.03	–
C27.0		IIa	1.93	3,140	0.93	1.07	–	0.39
C30.7	EF-T_s[f]	Ic	1.31	1,880	0.69	0.85	1.16	1.62
C31.6			3.49	4,860				
C34.3		Ib	2.63	3,370	1.04	1.06	–	0.91
C39.3		Ia	1.12	1,250	1.42	1.17	0.52	–
C40.3		Ib	5.50	6,000	0.92	1.14	0.87	0.94
C44.0		IIa	2.81	2,810	0.62	0.71	–	0.81
C44.6		IIa	4.41	4,350	0.61	0.80	2.03	0.91
C48.7		IIb	1.44	1,300	39.0	0.83	1.39	–
C55.0		Ic	1.38	1,100	0.45	0.80	1.67	–
C56.0		IIa	0.64	500	0.53	0.83	0.87	–
C58.5		IIa	0.47	350	0.63	0.90	0.95	–
C60.7		Ic	1.39	1,010	0.47	0.89	1.67	–
C61.0		Ic	1.88	1,360	0.63	0.89	1.47	–
C62.5		Ic	1.80	1,270	0.69	0.94	1.04	–
C62.7		Ic	2.57	1,800	0.63	0.73	2.09	–
C70.0		Ic	0.96	600	0.61	0.75	2.23	–
C78.0		Ic	3.07	1,730	0.79	1.01	1.29	–
C137		IIa	2.78	890	0.59	0.91	0.74	0.37
D15.2		Ic	1.01	2,920	0.53	0.78	1.50	–
D31.5		IIb	0.97	1,360	1.17	1.07	–	1.40
D32.5		IIa	1.39	1,880	0.77	0.92	–	0.89
D40.7		Ia	4.55	4,920	5.90	1.94	–	0.23
D44.5		IIa	6.94	6,860	0.56	1.35	0.48	–
D58.5	LysS	Ic	1.07	800	0.58	0.77	–	1.38
D74.0		Ic	3.28	1,950	0.51	0.68	–	1.52
D84.0	EF-G	Ic	16.6	8,700	0.60	0.88	1.43	1.70
D87.5		Ib	0.13	60	0.92	1.07	1.07	–
D94.0	PheS, β	Ic	2.05	960	0.54	0.94	1.12	1.10
D99.0		Ia	2.85	1,270	6.45	1.80	0.64	0.37
D100	LeuS	Ic	1.18	520	0.61	0.93	1.38	–
D102		Ia	0.81	350	5.10	1.61	0.78	–
D157	β	Ic	5.02	1,410	0.78	0.91	1.43	1.48
E25.4		IIa	1.02	1,770	0.68	0.45	–	0.10
E38.5		IIa	7.90	9,030	0.83	0.73	0.21	0.08
E39.8		Ia	1.68	1,860	8.72	2.90	0.53	–
E42.0	EF-Tu	Ic	55.5	58,140	0.71	0.97	–	1.53
E43.8		IIa	3.17	3,180	0.80	0.93	0.63	–
E58.0	ArgS	Ic	0.82	620	0.57	0.84	1.28	1.67
E77.5	GlyS	Ic	1.65	940	0.84	1.07	1.24	1.66
E79.0		IIb	1.58	880	1.28	1.25	1.66	2.97

Table 1—*Continued*

Protein Number[a]	Protein Identification[b]	Regulation Group[c]	Weight Fraction of Total Protein in Glucose Medium ($\alpha' \times 10^3$)	Number of Molecules per Genome in Glucose Medium[d]	Level in Each Medium Relative to the Level in Glucose Medium[e]			
					Acetate (k = 0.38)	Glycerol (k = 0.77)	Rich-Met (k = 1.50)	Rich (k = 1.98)
E106	ValS	Ic	1.39	580	0.58	0.82	1.32	1.52
E133		IIa	4.52	1,500	0.72	0.88	0.33	0.19
E140		IIa	2.70	850	0.80	0.89	0.21	0.41
F14.3		Ia	2.34	7,200	1.34	1.26	0.88	–
F14.7		Ia	3.13	9,370	1.24	1.23	1.01	–
F24.5		Ib	20.5	36,820	1.03	0.90	–	0.79
F28.7		Ic	1.19	1,820	0.45	0.93	–	1.03
F28.8		IIa	3.54	5,410	0.59	0.74	0.40	0.35
F29.7		Ic	1.19	1,760	0.63	0.98	1.07	–
F30.2		Ic	2.68	3,900	0.72	0.92	1.05	1.50
F32.3		Ia	2.53	3,450	3.96	2.01	0.38	0.19
F32.5		IIa	5.56	7,530	0.47	0.92	0.88	0.37
F36.9		Ib	0.69	820	0.84	0.98	–	0.93
F37.8		Ib	0.61	710	0.85	0.98	–	0.85
F38.0		IIa	9.49	10,990	0.56	0.77	–	0.31
F39.9		IIa	2.02	2,230	1.15	1.50	–	0.31
F41.8		Ic	1.40	1,470	0.64	0.80	–	1.29
F48.1	GluS, β	Ic	0.96	880	0.85	1.00	–	1.57
F50.3		Ia	1.73	1,510	4.32	2.02	0.88	0.14
F54.0		Ic	0.62	500	0.83	0.92	1.28	–
F54.4		IIa	3.50	2,830	0.77	0.87	0.15	0.17
F55.3		Ia	0.29	230	44.3	1.58	0.58	–
F56.0		Ia	1.21	950	5.53	3.30	1.07	–
F56.2		Ic	2.36	1,850	0.55	0.74	–	2.13
F58.5		Ic	1.36	1,020	0.78	0.96	1.02	–
F60.3		Ia	1.10	800	5.86	1.95	0.54	0.21
F63.4		Ic	0.40	280	1.03	1.05	–	1.40
F63.8		Ia	1.82	1,260	1.27	1.09	1.12	0.98
F64.5		IIa	3.16	2,160	0.88	1.00	1.29	0.92
F66.0		Ic	1.75	1,170	0.46	0.81	–	1.13
F82.5		Ia	0.48	260	3.63	2.42	0.76	0.39
F84.0		IIa	4.39	2,300	0.42	0.79	0.35	0.18
F84.1		Ia	0.88	460	1.27	1.05	0.79	0.83
F88.0		IIa	56.7	28,350	0.36	0.75	0.48	0.02
F99.0		Ic	8.54	3,800	0.62	0.80	2.34	2.87
F107	IleS	Ic	2.45	1,010	0.71	0.85	1.19	1.14
F178		Ib	0.99	240	1.06	0.94	1.02	–
G25.3		Ib	1.76	3,060	1.00	1.03	–	1.14
G27.2		Ia	0.89	1,440	12.78	7.50	0.70	–
G30.5		Ib	0.99	1,430	1.16	1.10	–	1.15
G31.6		IIb	0.25	350	2.05	0.80	–	1.23
G32.8		IIa	5.36	7,190	1.14	0.95	1.23	0.14
G36.0	PheS, α	Ic	1.07	1,310	0.53	0.87	–	1.05
G38.7		Ib	0.56	640	0.78	0.88	–	0.74
G40.3		IIa	1.64	1,790	0.56	0.72	–	0.36
G41.2		Ib	1.90	2,030	0.81	1.12	0.67	1.09
G41.3		Ic	1.09	1,160	0.77	1.04	–	5.04
G41.4		IIa	2.07	2,200	0.51	0.77	–	0.20
G42.0		Ic	1.83	1,920	0.80	0.94	–	1.88
G42.5		Ia	1.53	1,580	1.26	0.92	0.24	0.05
G43.2		IIa	6.75	6,880	0.41	0.55	0.47	0.15
G43.5		IIa	1.24	1,250	0.75	1.71	–	0.63
G43.8		IIa	7.97	8,010	0.80	0.60	0.66	0.25
G49.2		IIa	4.41	3,940	0.71	1.21	0.14	0.08
G50.5		IIb	4.60	4,010	1.96	1.24	1.37	1.41
G51.0		Ia	6.90	5,950	1.22	1.26	0.77	0.87
G51.1		IIa	0.64	550	1.09	1.26	–	0.39
G54.0		Ia	3.28	2,670	2.00	1.47	–	0.27
G54.6		Ia	0.65	520	3.75	2.12	–	0.30
G54.7		Ic	2.29	1,840	0.35	0.76	1.26	1.58
G57.0		Ia	0.55	420	1.49	1.70	1.00	0.99
G61.0	GlnS	Ic	1.14	820	0.70	0.99	–	1.31
G62.8		Ia	0.50	350	1.20	0.96	–	0.76

Table 1—*Continued*

Protein Number[a]	Protein Identification[b]	Regulation Group[c]	Weight Fraction of Total Protein in Glucose Medium ($\alpha' \times 10^3$)	Number of Molecules per Genome in Glucose Medium[d]	Level in Each Medium Relative to the Level in Glucose Medium[e]			
					Acetate (k = 0.38)	Glycerol (k = 0.77)	Rich-Met (k = 1.50)	Rich (k = 1.98)
G63.0		Ic	0.64	450	0.65	0.96	–	1.35
G65.0	ThrS	Ic	0.86	580	0.72	0.85	–	1.67
G70.0		Ic	0.71	450	0.56	1.31	–	1.26
G71.5		Ib	0.55	340	0.73	1.00	–	0.90
G72.5		Ia	0.27	160	0.96	0.83	–	0.65
G74.0		Ia	3.25	1,930	2.54	1.72	–	0.72
G76.0		Ia	0.17	100	9.05	1.30	0.85	–
G78.0		Ic	0.85	480	0.58	0.98	–	1.03
G93.0		Ia	2.18	1,030	1.52	1.33	0.95	0.74
G97.0		Ia	1.99	900	4.45	1.06	0.86	0.30
G117		Ic	1.75	660	0.78	0.92	1.07	–
G127		Ib	0.25	90	0.83	0.88	0.93	–
H35.0		IIa	2.98	3,750	0.55	0.80	–	0.30
H47.4		Ia	1.12	1,040	4.37	1.68	–	0.07
H52.7		Ic	1.69	1,410	0.35	0.60	1.84	2.41
H54.6		Ia	0.65	520	4.07	2.35	0.67	0.82

[a] Proteins are numbered as described in the Appendix. Letter prefixes refer to the position of the protein in the axis parallel to the isoelectric focusing dimension (A–H; acid to base). Numbers refer to the apparent molecular weight of the protein as determined by the distance moved in the SDS electrophoresis dimension (56.5; molecular weight 56,500).

[b] Abbreviations used: (S) aminoacyl tRNA synthetase; (α) 39,000 molecular weight subunit of RNA polymerase; (β) 155,000 molecular weight subunit of RNA polymerase; (L7, L12, S1, S6) ribosomal polypeptides; (EF-Ts, EF-G, EF-Tu) elongation factors; (A) A protein (Subramanian et al., 1976).

[c] The proteins have been grouped according to the way in which their relative levels change in cells grown in different media. As the growth rates increase, the relative levels of Ia proteins decrease, Ib proteins do not change, Ic proteins increase, IIa have a maximum and IIb have a minimum.

[d] The number of polypeptide molecules per genome equivalent DNA for cells growing in glucose-minimal medium was calculated as follows: $\frac{\alpha'}{MW} \times 110$ daltons per amino acid $\times$ 4 $\times$ 10^8 amino acids per genome. To calculate the number of molecules per genome for growth in the two rich media, multiply the above product by the relative level of the protein. Since the amino acids per genome are different at generation times greater than 40 min, 2.3 $\times$ 10^8 and 3.5 $\times$ 10^8 must be substituted for acetate and glycerol media, respectively.

[e] Proteins for which a value is shown for glucose-rich medium lacking methionine (Rich-Met) and not for glucose-rich medium (Rich) are those which were measured only in the experiments using ^3H and ^{35}S. Proteins for which a value is shown for Rich and not for Rich-Met were measured only in the experiments using ^3H and ^{14}C. Where a value is shown for both of these media, the same protein was measured by each method, and in these cases, the values for acetate- and glycerol-minimal media are averages of the results in the two sets of experiments.

[f] Elongation factor Ts forms a double spot, C30.7 and C31.6. Separate α' values were measured in glucose medium, but the relative levels in different media were measured on the combined spots.

of the ampholytes used here, gives α' values ranging 0.57–0.68 (Figure 1). The regulation of the major part of the cell's protein mass has therefore been characterized at these growth rates.

A broad range of gene expression is represented among the 140 proteins. The protein which was found to be the most abundant is EF-Tu, which in rich medium is synthesized at a rate of about 49 molecules per genome per second. As found previously (Furano, 1975; Pedersen et al., 1976a), this is about 5 times as fast as the rate of synthesis of ribosomal proteins. EF-Tu is the product of two genes, *tufA* and *tufB* (Jaskunas et al., 1975). Estimates of the relative expression from the two genes (Pedersen et al., 1976a, 1976b) show that the *tufA* gene is the most active gene in E. coli, synthesizing about 34 molecules per genome per second in the rich medium. Ribosomal proteins S1 and S6, and EF-G are found to be made at about 10 molecules per genome per second, and ribosomal protein (L7 + L12) is synthesized at 2–3 times that rate, making the *rplL* gene the second most active gene. A single gene is responsible for this synthesis, since one mutation gives an electrophoretic mobility change of both L7 and L12 species (Watson et al., 1975). No previous measurement of the total amount of S1, which might behave as an initiation factor (Dahlberg and Dahlberg, 1975), has been made. At steady state growth conditions, it behaves as a ribosomal protein, but during amino acid starvation in *relA* strains, it is regulated differently from typical ribosomal proteins (Reeh, Pedersen and Friesen, 1976). The A protein (Subramanian, Haase and Giessen, 1976) B56.5, which in stationary phase cells is found in association with ribosomes, is also made at a rate close to that of ribosomal proteins. It has strong and unusual regulation, quite different from ribosomal proteins under non-steady state growth conditions (Reeh et al., 1976). Besides the ribosomal proteins and

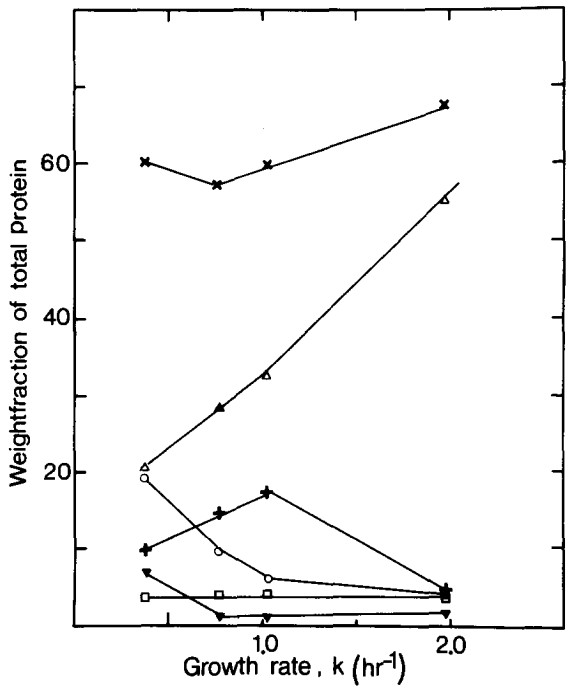

Figure 1. Variation of Protein Amounts in Different Regulatory Groups as a Function of Growth Rate in E. coli NC3

The ordinate values are the sum of the weight fractions (α; calculated from the data of Table 1) of the individual proteins in each regulation group in cells grown in each medium. For each group, only one value is given for rich medium. Values have been extrapolated from rich-methionine medium where only this value has been determined. The resulting error is negligible since these proteins constitute only about 10% by weight of the total. To group Ic have been added the weight amounts (13%) of ribosomal proteins (Dennis and Bremer, 1974) corrected for the ribosomal protein directly measured in this study (S_1, S_6, 4 copies per ribosome of L7/L12). (○) group Ia (33 proteins); (□) group Ib (16 proteins); (△) group Ic (105 proteins); (+) group IIa (32 proteins); (▼) group IIb (5 proteins); (×) total recovered protein.

translation factors, high synthesis rates are found also for protein F24.5 (16 molecules per genome per second), for F88.0 (8 molecules per genome per second) and for C48.7 (3 molecules per genome per second) in one or another of the five media. The weakest promotor of the proteins investigated is that of protein G127, which is made at a rate of 1 per 3.7 min.

Since the range of α' is about 100 fold in each of the regulatory groups, promotors of varying strength can be subject to any of the regulatory patterns by which we have organized our data.

The regulation of the cell's proteins was studied to investigate the possibility that large numbers of proteins might be regulated by a common molecular mechanism—for example, by the concentration of an effector molecule. Do we see any evidence for this? A glance at the regulation of individual proteins within each group shows that at least some groups are quite heterogeneous, and in all groups there are some proteins regulated differently from the bulk of the proteins within the group. Within group Ia, however, a hyperbola-shaped subgroup can be seen; group Ib by definition is a homogeneous group, with 16 proteins exhibiting the same shape of regulation curve; and within group Ic, there is evident one subgroup, Ic-1, with regulation like ribosomal proteins, and another subgroup, Ic-2, which shows a shallower slope of relative level as a function of growth rate than do ribosomal proteins. Groups IIa and IIb seem to be heterogeneous, and clear subgroups might become evident as more proteins are characterized.

Some genes might be devoid of any specific regulation mechanism. Their expression should then vary passively by competition for—for example—a transcription component, the availability of which should depend upon the transcription from the actively regulated genes (Maaløe, 1969). From the point of view of economy, one would expect the most active controls to be evolved for the most abundant proteins. The data in Table 1 agree with this expectation. Conversely, genes without active regulation should be found among those with the lowest levels of expression. Examining the α' values at all growth rates (calculated from Table 1) shows that 14 proteins in one or more media have α' values lower than 0.3×10^{-3}. Among these 14, nine proteins (67%) are in group Ia, which contains only 20% of the number of proteins investigated. Of these nine proteins, eight have a regulation curve closely resembling a hyperbola. (Deviation from the theoretical curve, 1/k, might be caused by variations in gene dosage.) This clearly is an overrepresentation of group Ia among the weakest promotors. The regulatory pattern of group Ia is what would be observed if these weak promotors were unregulated and were saturated with all components of the cell's transcription apparatus within the span of growth rates studied here. One regulatory mechanism in common for some proteins within group Ia might therefore by the *absence* of active regulation.

Within group Ib, a unique regulation is found for protein F24.5 under amino acid starvation (Reeh et al., 1976). Protein F24.5 is the SSP protein of Reeh et al. (1976), which constitutes the dominant part of the residual protein synthesis in *relA*$^+$ strains under severe amino acid starvation. Thus the constant level of group Ib proteins during exponential growth is not brought about by a single common regulatory mechanism.

Within group Ic, the subgroup Ic-1, regulated as ribosomal proteins, might have a common regulatory mechamism. Under special cases such as amino acid starvation (Blumenthal et al., 1976a; Reeh et al., 1976), however, at least S1, EF-G, and

valyl and arginyl-tRNA synthetases are regulated differently from the ribosomal proteins, which under this condition are regulated as a homogeneous group (Dennis and Nomura, 1974). Some of the proteins in group Ic-2—for example, RNA polymerase subunit α, EF-Tu and isoleucyl-tRNA synthetase—are regulated differently from each other during amino acid starvation (Blumenthal et al., 1976a; Reeh et al., 1976).

In summary, we have found little evidence that the proteins within a group share a common regulatory mechanism; these may exist, but if so their effect is obscured by specific control mechanisms.

Few of the 140 proteins studied here have been identified, but it will be useful to comment on the nature of the proteins in each group.

Group Ia

No proteins in this group have been identified. Catabolic enzymes, including β-galactosidase (Wanner, Kodaira and Neidhardt, 1977), which are subject to cAMP control exhibit the behavior of this group. Enzymes of oxidative metabolism, including Krebs cycle enzymes and cytochromes, should also exhibit this pattern (Magasanik, 1961). Some proteins specific for acetate utilization are also temporarily placed in this group, but are repressed in other carbon sources giving a slower growth rate (S. Pedersen, unpublished observation).

Group Ib

No proteins of this group have been identified in this study, but the UV5 promotor of the *lac* operon results in the synthesis of β-galactosidase in this pattern (Wanner et al., 1977), and the transcription factors ρ (Blumenthal, Reeh and Pedersen, 1976b) and σ (Iwakura and Ishihama, 1976) also appear relatively constant in level.

Group Ic

This group, together with about 50 unrecovered ribosomal proteins, contains 105 proteins. Most are already identified as ribosomal proteins, elongation factors, aminoacyl tRNA synthetases and subunits of RNA polymerase. It seems probable that many of the remaining proteins in this group are also concerned with translation and transcription, and perhaps replication. The increase of group Ic occurs at the expense of proteins of one regulation group (Ia) from acetate to glucose-minimal medium and at the expense of proteins of another regulation group (IIa) from glucose-minimal to rich medium.

Group IIa

These enzymes are most probably concerned with the biosynthesis of amino acids, nucleotides and other building blocks of protoplasm. None has been identified, but the striking effect of rich medium on enzymes of this sort is a common observation (Camakaris and Pittard, 1971; Rose and Yanofsky, 1972; Cortese et al., 1974; Stephens, Artz and Ames, 1975). Some of these proteins, particularly F88.0, are extremely abundant (28,000 molecules per genome in glucose). Protein F88.0 is synthesized at about the same rate per second in glucose and in rich medium minus methionine, and that rate might very well represent the maximal activity of its promotor, leading to a relative decrease going from glucose to rich medium lacking methionine. Addition of methionine reduces that rate at least 50 fold, suggesting that F88.0 is a biosynthetic enzyme for methionine, although its extreme abundancy is a puzzle.

Group IIb

This group constitutes a minor fraction of the cell's protein. Protein C48.7 seems specific for acetate utilization, and shows a barely significant minimum level in glycerol medium. The remaining proteins constitute about 1% of total cell protein.

Further work will include exploration of the hypothesis of the nature of the proteins in these groups, and will address the question of the nature of the controls on these proteins, giving particular attention to controls that might operate on all or most proteins of an individual group (Maaløe, 1969; Rose and Yanofsky, 1972; Rickenberg, 1974; Stephens et al., 1975).

Experimental Procedures

The E. coli B/r derivative NC3 (Neidhardt et al., 1977) was grown in five media which contained MOPS medium (Neidhardt, Bloch and Smith, 1974) supplemented with either acetate 0.4%, glycerol 0.4%, glucose 0.4%, rich-methionine medium and rich medium. Rich medium contained 20 amino acids, four purines and pyrimidines, and five vitamins as previously described (Neidhardt et al., 1977). Cells were grown aerobically with rotary shaking at 37°C. Growth was monitored at 420 nm, and growth rates are expressed as the specific growth rate constant k = ln 2/mass doubling time (hr). For the five media listed above, the value of k was 0.38, 0.77, 1.03, 1.50 and 1.98 hr^{-1}, respectively.

Cultures were labeled for two to three generations in media containing either ^{14}C-glucose, ^{3}H-leucine, ^{14}C-leucine, ^{3}H-isoleucine + ^{3}H-leucine or $^{35}SO_4$. Details of the labeling methods have been described (Neidhardt et al., 1977). The preparation of cell extracts, electrophoresis, staining and counting procedures have been described (Pedersen et al., 1976a; Neidhardt et al., 1977).

Identification of proteins and the nomenclature of spots on these gels are described in the Appendix.

Acknowledgments

We thank Mr. Alan Baumgarten for expert assistance in tabulating and processing the data. P. O'Farrell very generously communicated to us the elements of his method prior to publication. This work was supported by grants from the NSF and the National Institute of General Medical Sciences. S. P. was supported by

Statens Naturvidenskabelige Forskningsråd, Denmark, and by a NATO postdoctoral fellowship.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 13, 1977; revised January 31, 1978

References

Blumenthal, R. M., Reeh, S. V. and Pedersen, S. (1976b). Regulation of the transcription factor ρ and α subunit of RNA polymerase in *Escherichia coli* B/r. Proc. Nat. Acad. Sci. USA *73*, 2285–2288.

Blumenthal, R. M., Lemaux, P. G., Neidhardt, F. C. and Dennis, P. P. (1976a). The effects of the *relA* gene on the synthesis of amino-acyl-tRNA synthetases and other transcription and translation proteins in *Escherichia coli* B. Mol. Gen. Genet. *149*, 291–296.

Camakaris, J. and Pittard, J. (1971). Repression of 3-deoxy-D-arabino-heptulosonic acid-7-phosphate synthetase (trp) and enzymes of the tryptophan pathway in *Escherichia coli*. J. Bacteriol. *107*, 406–414.

Cortese, R., Landsberg, R., von der Haar, R. A., Umbarger, H. E. and Ames, B. N. (1974). Pleiotrophy of *hisT* mutants blocked in pseudo-uridine synthesis in tRNA: leucine and isoleucine-valine operons. Proc. Nat. Acad. Sci. USA *71*, 1857–1861.

Dahlberg, A. D. and Dahlberg, J. E. (1975). Binding of ribosomal protein S1 of *Escherichia coli* to the 3′ end of 16S rRNA. Proc. Nat. Acad. Sci. USA *72*, 2940–2944.

Dennis, P. P. (1974). *In vivo* stability, maturation and relative differential synthesis rates of individual ribosomal proteins in *Escherichia coli*. J. Mol. Biol. *88*, 24–41.

Dennis, P. P. and Bremer, H. (1974). Macromolecular composition during steady growth of *Escherichia coli* B/r. J. Bacteriol. *119*, 270–281.

Dennis, P. P. and Nomura, M. (1974). Stringent control of ribosomal protein gene expression in *Escherichia coli*. Proc. Nat. Acad. Sci. USA *71*, 3819–3823.

Deusser, E. (1972). Heterogeneity of ribosomal populations in *Escherichia coli* cells grown in different media. Mol. Gen. Genet. *119*, 249–258.

Furano, A. V. (1975). Content of elongation factor Tu in *Escherichia coli*. Proc. Nat. Acad. Sci. USA *72*, 4778–4784.

Gordon, J. (1970). Regulation of the *in vivo* synthesis of the polypeptide chain elongation factors in *Escherichia coli*. Biochemistry *9*, 912–917.

Iwakura, Y. and Ishihama, A. (1975). Biosynthesis of RNA polymerase in *Escherichia coli*. II. Control of RNA polymerase synthesis during nutritional shift up and down. Mol. Gen. Genet. *142*, 67–84.

Jaskunas, S. R., Lindahl, L., Nomura, M. and Burgess, R. R. (1975). Identification of two copies of the gene for the elongation factor EF-Tu in *E. coli*. Nature *257*, 458–462.

Kaltschmidt, E. and Wittmann, H. G. (1970). Ribosomal proteins. VII. Two-dimensional polyacrylamide gel electrophoresis for finger printing of ribosomal proteins. Anal. Biochem. *36*, 401–412.

Maaløe, O. (1969). An analysis of bacterial growth. Dev. Biol. Suppl. *3*, 33–58.

Magasanik, B. (1961). Catabolite repression. Cold Spring Harbor Symp. Quant. Biol. *26*, 249–254.

Matzura, H., Hansen, B. S. and Zeuthen, J. (1973). Biosynthesis of the β and β' subunits of RNA polymerase in *Escherichia coli*. J. Mol. Biol. *74*, 9–20.

Moses, V. and Wild, D. G. (1969). Soluble protein profiles in *Escherichia coli*. Folia Microbiologica *14*, 305–309.

Neidhardt, F. C., Bloch, P. L. and Smith, D. F. (1974). Culture medium for enterobacteria. J. Bacteriol. *109*, 736–747.

Neidhardt, F. C., Bloch, P. L., Pedersen, S. and Reeh, S. (1977). Chemical measurement of steady state levels of ten aminoacyl-transfer ribonucleic acid synthetases in *Escherichia coli*. J. Bacteriol. *129*, 378–387.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. *250*, 4007–4021.

Pedersen, S., Reeh, S. R., Parker, J., Watson, R. J., Friesen, J. D. and Fiil, N. P. (1976a). Analysis of the proteins synthesized in ultraviolet light-irradiated *Escherichia coli* following infection with the bacteriophage λdfus-3 and λdfus-3. Gen. Genet. *144*, 339–344.

Pedersen, S., Blumenthal, R. M., Reeh, S., Parker, J., Lemaux, P., Laursen, R. A., Nagarkatti, S. and Friesen, J. F. (1976b). A mutant of *Escherichia coli* with an altered elongation factor Tu. Proc. Nat. Acad. Sci. USA *73*, 1698–1701.

Raunio, R. and Sarimo, S. (1965). Changes in protein composition of *Escherichia coli* during the active growth phase. Acta Chem. Scand. *19*, 2428–2435.

Reeh, S., Pedersen, S. and Friesen, J. D. (1976). Biosynthetic regulation of individual proteins in *relA*$^+$ and *relA* strains of *Escherichia coli* during amino acid starvation. Mol. Gen. Genet. *149*, 279–289.

Rickenberg, H. V. (1974). Cyclic AMP in prokaryotes. Ann. Rev. Microbiol. *28*, 353–369.

Rose, J. K. and Yanofsky, C. (1972). Metabolic regulation of the tryptophan operon of *Escherichia coli:* repression-independent regulation of transcription initiation frequency. J. Mol. Biol. *69*, 103–118.

Schleif, R. (1967). Control of production of ribosomal protein. J. Mol. Biol. *27*, 41–55.

Stephens, J. C., Artz, S. W. and Ames, B. N. (1975). Guanosine d′-disphosphate 3′-disphosphate (ppGpp): positive effector of histidine operon transcription and general signal for amino acid deficiency. Proc. Nat. Acad. Sci. USA *72*, 4389–4393.

Subramanian, A. R., Haase, C. and Geisen, M. (1976). Isolation and characterization of a growth-cycle-reflecting, high-molecular-weight protein, associated with *Escherichia coli* ribosomes. Eur. J. Biochem. *67*, 591–601.

Wanner, B. L., Kodaira, R. and Neidhardt, F. C. (1977). Physiological regulation of a decontrolled *lac* operon. J. Bacteriol. *130*, 212–222.

Watson, R. J., Parker, J., Fiil, N. P., Flaks, S. G. and Friesen, J. D. (1975). New chromosomal location for structural genes of ribosomal proteins. Proc. Nat. Acad. Sci. USA *72*, 2765–2769.

Appendix: A Nomenclature for E. coli Proteins

Steen Pedersen and Solvejg Reeh
Department of Molecular Biology
University of Aarhus
DK-8000 Aarhus C, Denmark
James D. Friesen
Biology Department
York University
Downsview, Ontario M3J 1P3, Canada
Philip L. Bloch and Frederick C. Neidhardt
Department of Microbiology
The University of Michigan
Ann Arbor, Michigan 48109

Considerable information about large numbers of the major proteins of E. coli will be generated by

many laboratories using the two-dimensional gel electrophoresis technique of O'Farrell (1975). To organize and communicate these data, it is useful to have a code identifying each protein as it appears in this separation system. We have been using such a code, and we describe it here.

In our code, a protein is designated by the two properties, isoelectric point and molecular weight, by which proteins are separated in the O'Farrell technique. Another useful feature is that the nomenclature is based on the positions of internal markers. Proteins that are synthesized under all growth conditions and are easily recognized in the spot pattern of the gels make excellent markers, and we therefore chose to use certain proteins involved in translation and transcription for this purpose.

The autoradiogram of the gel is positioned with the acidic proteins to the right and with the direction of development of the second dimension pointed downward. A set of lines perpendicular to the border between the stacking and running gel is now drawn on the autoradiogram. The lines divide the gel into zones named A, B . . . , I, the A zone being the most acidic. An example is shown in Figure 1 (right panel). The line defining the A/B border is drawn between ribosomal proteins L7 and L12, the B/C line through the center of RNA polymerase subunit α, the C/D line through the center of the EF-Ts spot, the D/E line through the center of the EF-G spot, the E/F line through the center of the EF-Tu spot, the F/G line through the center of the isoleucyl-tRNA synthetase spot, and the G/H line through the center of an unidentified protein spot indicated in Figure 1. Additional proteins can be resolved using an ampholine mixture of pH 3–10 (2%) as shown in Figure 1 (left panel) instead of the used mixture of pH 3–10 (0.4%) plus pH 5–7 (1.6%). Among these proteins is ribosomal protein S2. We suggest that the H/I line be drawn, as indicated, through S2, but we have not yet used this additional zone.

Several proteins have been identified in this gel system (see below). The molecular weight of these proteins, as determined by other investigators, is plotted in Figure 2 as a function of their migration

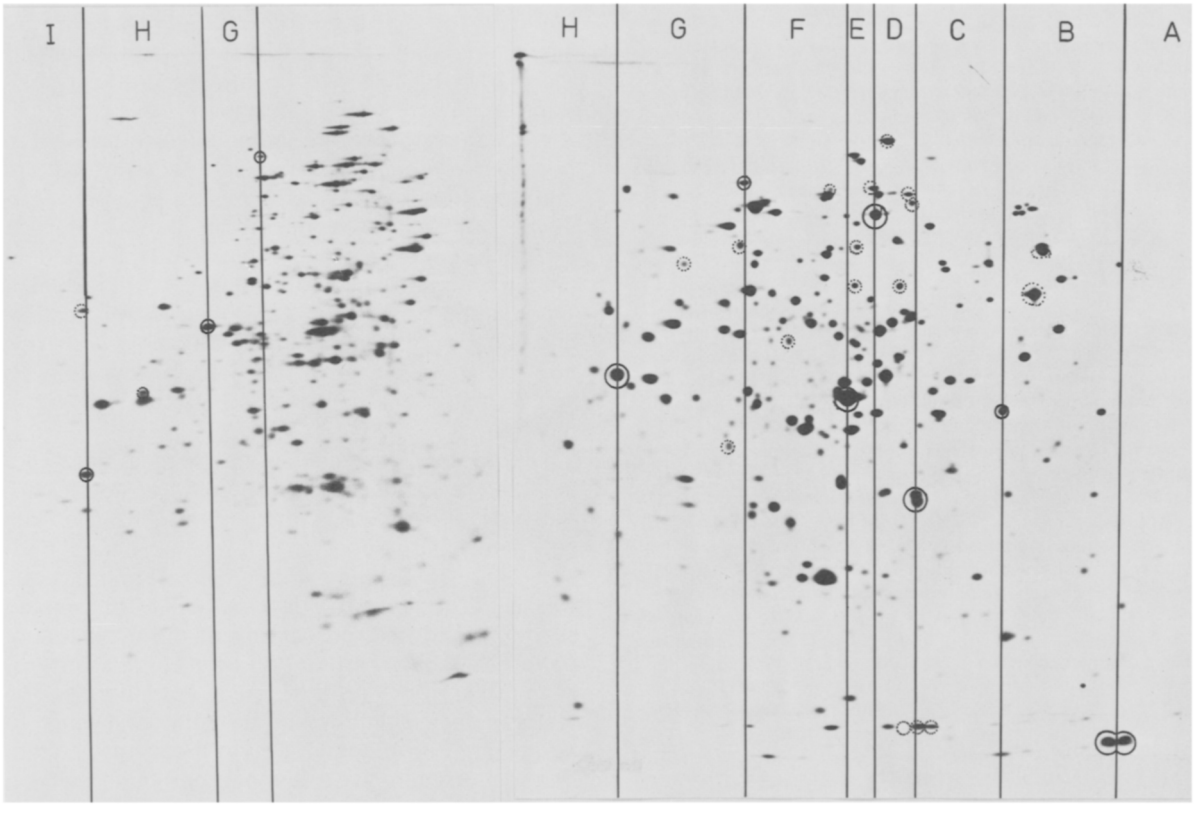

Figure 1. Two-Dimensional Gels of Extracts from E. coli B/r Strain NC3 Grown in Glucose-Minimal Medium and Labeled with ^{35}S-Methionine

(Left) gel using 2% pH 3–10 ampholine mixture in the first dimension. (Right) gel using 0.4% pH 3–10 + 1.6% pH 5–7 ampholine mixture in the first dimension. The vertical lines define the zones with different isoelectric points. The solid circles indicate the protein spots used for defining the zones; the dotted circles indicate the protein spots identified so far.

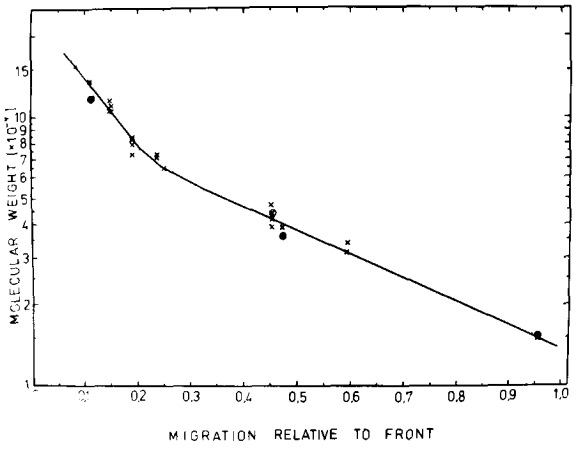

Figure 2. Migration of Several Identified Proteins in the Second Dimension as a Function of Their Molecular Weight

Closed circles, open circles and crosses represent molecular weights determined by the amino acid sequence, a partial sequence and by other means, respectively. With decreasing molecular weight (MW), the proteins are as follows: β subunit of RNA polymerase: MW 155,000 daltons (Burgess, 1969); β-galactosidase: MW 116,248 (Fowler and Zabin, 1977), 135,000 (Zabin and Fowler, 1970); isoleucyl-tRNA synthetase: MW 114,000 (Arndt and Berg, 1970), 110,000, 105,000 (Berthelot and Yaniv, 1970); valyl tRNA synthetase: MW 100,000, 105,000 (Berthelot and Yaniv, 1970); elongation factor G: MW 83,000, 80,000, 72,000 (Kaziro et al., 1969; Leder et al., 1969; Lengyel, 1974); arginyl-tRNA synthetase: MW 74,000, 72,000 (Hirschfield and Bloemers, 1969); ribosomal protein S1: MW 65,000 (Wittmann, 1974); elongation factor Tu: MW 47,000, 44,000, 39,000 (Lucas-Lenard, 1971; Lengyel, 1974; Wittinghofer and Leberman, 1976; Laursen et al., 1977; Nakamura et al., 1977); α subunit of RNA polymerase: MW 39,000 (Burgess, 1969), 36,512 (Ovchinnikov et al., 1977); elongation factor Ts: MW 34,000 (Lengyel, 1974), 31,500 (Wittinghofer and Leberman, 1976); ribosomal protein S6: MW 15,600 (Wittman, 1974), 15,706 (Hitz et al., 1977). For some proteins, like S6, there is excellent agreement with the molecular weight determined by SDS gel electrophoresis (Wittmann, 1974) and by the primary structure of the protein (Hitz et al., 1977). For other proteins—for example, β-galactosidase (Fowler and Zabin, 1977), the α subunit of RNA polymerase (Ovchinnikov et al., 1977) and several ribosomal proteins (sequences reviewed by Wittmann-Liebold et al., 1977)—the apparent molecular weight in SDS gels is up to 30% higher than the one calculated from the primary sequence. The groups working on the primary structure of EF-Tu estimate the molecular weight to be 44,000 (Laursen et al., 1977) and 47,000 daltons (Nakamura et al., 1977), so it seems that the apparent molecular weight determined by SDS gel electrophoresis underestimates the molecular weight of Ef-Tu by some 5–10%. Given such small, and for many purposes insignificant, variations between apparent molecular weight in SDS and true molecular weight, we believe that the line given in Figure 2 is sufficient for defining the molecular weight solely for this nomenclature purpose.

relative to that of the front. All distances are measured from the border between the stacking and running gel of the second dimension. The best line through these marker proteins (Figure 2) is used to define the apparent molecular weight of each protein. The code designation of a given spot consists of a letter, indicating the zone of its isoelectric point, and three numbers, indicating the apparent molecular weight of the protein $\times 10^{-3}$. Protein C57.3, therefore, is a protein with an isoelectric point between that of the α subunit of RNA polymerase and EF-Ts, and with an apparent molecular weight of 57,300 daltons. The uncertainty of the molecular weight determination probably makes the third numeral meaningless, but it is useful to indicate the relative positions of proteins which have similar molecular weights.

Protein spots that are centered on one of the zonal division lines are assigned the letter designating the more acidic of the two zones. Should two proteins in the same zone appear to have exactly the same molecular weight, we suggest the use of the suffixes a, b, c . . . in order of their relative intensity on autoradiograms made from extracts of cells grown in glucose-minimal medium containing ^{35}S-methionine.

Figure 1 shows the location of identified proteins, including those used as markers; Table 1 gives their code designations together with the basis for each identification. Different strains of E. coli (B/r, the B strain AS19, K12 and C) have very similar spot patterns. Most proteins, and especially the marker proteins, have virtually identical positions on gels from these strains.

We suggest that the nomenclature presented here be adopted to minimize the confusion of multiple, unrelated systems of nomenclature. Because our system is based on the position of internal marker proteins, misidentifications and ambiguities among laboratories can be reduced. This coding procedure is simple and informative, and because reproducibility of the gels is high, this nomenclature can be used to focus attention on small areas of the gels for interlaboratory communication.

This system, however, is not the definitive nomenclature that should eventually be developed. As more experience is gained with the O'Farrell gels, many additional proteins will be identified, and technical improvements and simplifications will be introduced. The development of automated scanning and data processing of the protein spot patterns is proceeding in several laboratories. From this work will emerge a definition of the features desired in a final (international) system of nomenclature. At some point in the future, a standard gel of the total protein of a particular cell can be prepared under defined and easily reproduced conditions, and at that time, every spot will be named according to an agreed upon coordinate system.

Table 1. Identification of Polypeptides[a]

Polypeptide (Code Designation)	Migration with: Purified Polypeptide	One Polypeptide of Protein Mixture	One Polypeptide of Phage-Coded Protein	References
S1(B65.0)	–	Qβ replicase; 30S ribosome	–	Wahba et al. (1974)
S2(H27.0)	–	30S ribosome	λdpolC9	Friesen et al. (1976a)
S6(C14.8, C14.9, D14.7)[b]	–	30S ribosome	–	Lindahl et al. (1975); Pedersen et al. (1976)
L7(A13.0)	–	50S ribosome	$\lambda drif^{d}_{18}$	Watson et al. (1975); O'Farrell (1975)
L12(B13.0)	–	50S ribosome	$\lambda drif^{d}_{18}$	O'Farrell (1975)
EF-Tu(E42.0)	+	EF-T, Qβ replicase	$\lambda drif^{d}_{18}$, λdfus3	Jaskunas et al. (1975); Pedersen et al. (1976)
EF-Ts(C30.7, C31.6)[c]	+	EF-T, Qβ replicase	λdpolC9	Friesen et al. (1976a)
EF-G (D83.0)	+	–	λdfus3	Jaskunas et al. (1975)
RNA Polymerase[d]				Kirschbaum and Scaife (1974); Jaskunas et al. (1975);
α Subunit (B40.7)	–	RNA polymerase	λdfus3	Friesen et al. (1976a)
β Subunit (D157)	–	RNA polymerase	$\lambda drif^{d}_{18}$	Same
β Subunit (I160)	–	RNA polymerase	$\lambda drif^{d}_{18}$	Same
Stringent Factor (H78.0)[e]	+	–	λdpyrG	Friesen et al. (1976b)
β-Galactosidase (E123)[f]	–	–	–	
IF2B(F106)	–	IF2	–	Lengyel (1974)
ρ Factor[g]	+	–	λdilv	Blumenthal et al. (1976)
A Protein (B56.5)	+	–	–	Subramanian et al. (1976)
Ten Aminoacyl-tRNA Synthetases[h]	+	–	–	Neidhardt et al. (1977)

[a] Specific polypeptides were located by their migration with markers. The marker polypeptide was pure—one of the subunits of a pure protein or of a purified protein complex, or one of a mixture of a small number of proteins produced by ultraviolet-irradiated cells infected with λ phage bearing small known portions of bacterial chromosome.

[b] S6 is known to contain multiple forms shown to be interconvertible on the two-dimensional gel (S. R. Reeh and S. Pedersen, manuscript in preparation). A mutant containing only one form gave only protein spot D14.7 (Hitz et al., 1977).

[c] The two spots are probably derived from each other in an artifactual process. Pulse-chase experiments do not show a precursor-product relation.

[d] The α subunit was distinguished from the other two subunits by its molecular weight. β and β′ were distinguished on the basis of known pI (Fujiki and Zurek, 1975). The σ subunit is not yet identified, but is one of several spots in the region of B90.

[e] In addition, H78.0 has been shown to be electrophorectically altered in several *relA* mutants (S. Pedersen, unpublished results).

[f] β-Galactosidase has been identified by the fact that protein spot E123 is inducible by IPTG and by the fact that it is present in the extract of a constitutive strain.

[g] Termination factors ρ from E. coli K12 and E. coli B migrate with spot H51.3 and I48.8, respectively. λdilv was isolated and shown to transduce the *rho* gene (P. Jorgensen, unpublished results).

[h] Ten aminoacyl-tRNA synthetases identified by migration with purified proteins (Neidhardt et al., 1977). The isoleucyl-tRNA synthetase identification was corroborated with λddapB2 (Friesen et al., 1976c).

Acknowledgments

The purified proteins were obtained from many individuals, to whom we are very grateful. Purified 30S and 50S ribosomal proteins: Dr. L. Lindahl; Qβ replicase: Dr. T. Blumenthal; RNA polymerase: Drs. F. Engbaek and H. Weissbach. Transcription termination factor rho from E. coli K12: Dr. F. Engbaek; from E. coli: Dr. J. Richardson. Elongation factors Tu: Drs. T. Furano, C. Ganoza and H. Weissbach. EFTs: Drs. H. Weissbach and C. Ganoza. EFG: Drs. H. Weissbach and J. Gordon. IF 1,2,3: Drs. H. Weissbach and J. Ravel. Stringent factor: Dr. F. S. Pedersen. A protein: Dr. A. Subramanian. Arginyl-tRNA synthetase, Arg RS: Dr. W. Konigsberg, Ile RS: Drs. M. Yarus and W. Konigsberg. Glu RS: Dr. H. Weissbach. Gly RS: Dr. P. Berg. Leu RS: Dr. H. Weissbach. Lys RS: Dr. I. Hirshfield. Phe RS: Dr. A. Bock. Thr RS: Dr. H. Weissbach. ValRS: Dr. D. Soll. Trp RS: Dr. K. Muench.

S.P. and S.R. were supported by grants from NATO, the Danish Science Research Council and the Carlsberg Foundation. J.D.F. was supported by grants from the National Research Council of Canada and the National Cancer Institute of Canada. F.C.N. and P.L.B. received support from the NSF and the National Institute of General Medical Sciences.

References

Arndt, D. J. and Berg, P. (1970). Isoleucyl transfer ribonucleic acid synthetase is a single polypeptide chain. J. Biol. Chem. *245*, 665–667.

Berthelot, F. and Yaniv, M. (1970). Presence of one polypeptide chain in valyl and isoleucyl tRNA synthetases from *Escherichia coli*. Eur. J. Biochem. *16*, 123–125.

Blumenthal, R. M., Reeh, S. V. and Pedersen, S. (1976). Regulation of transcription factor ρ and the α subunit of RNA polymerase in *Escherichia coli* B/r. Proc. Nat. Acad. Sci. USA *73*, 2285–2288.

Burgess, R. R. (1969). Separation and characterization of the subunits of ribonucleic acid polymerase. J. Biol. Chem. *244*, 6168–6176.

Fowler, A. V. and Zabin, I. (1977). The amino acid sequence of β-galactosidase of *Escherichia coli*. Proc. Nat. Acad. Sci. USA *74*, 1507-1510.

Friesen, J. D., Parker, J., Watson, R. J., Bendiak, D., Reeh, S., Pedersen, S. and Fiil, N. P. (1976a). A transducing bacteriophage carrying the structural gene for elongation factor Ts. Mol. Gen. Genet. *148*, 93-98.

Friesen, J. D., Parker, J., Watson, R. J., Fiil, N. P., Pedersen, S. and Pedersen, F. S. (1976b). Isolation of a lambda transducing bacteriophage carrying the *relA* gene of *Escherichia coli*. J. Bacteriol. *127*, 917-922.

Friesen, J. D., Parker, J., Watson, R. J., Fiil, N. P. and Pedersen, S. (1976c). Isolation of a transducing phage carrying *rpsT*, the structural gene for ribosome protein S20. Mol. Gen. Genet. *144*, 115-118.

Fujiki, H. and Zurek, G. (1975). The subunits of DNA-dependent RNA polymerase from *E. coli:* amino acid analysis and primary structure of the N-terminal regions. FEBS Letters *55*, 242-244.

Hirschfield, I. N. and Bloemers, H. P. J. (1969). The biochemical characterization of two mutant arginyl tRNA synthetases from *Escherichia coli* K12. J. Biol. Chem. *244*, 2911-2916.

Hitz, H., Schafer, D. and Wittmann-Liebold, B. (1977). Determination of the complete amino acid sequence of protein S6 from the wildtype and a mutant of *Escherichia coli*. Eur. J. Biochem. *75*, 497-512.

Jaskunas, S. R., Lindahl, L., Nomura, M. and Burgess, R. R. (1975). Identification of two copies of the gene for the elongation factor EF-Tu in *E. coli*. Nature *257*, 458-462.

Kaziro, Y., Inoue, N., Kuriki, Y., Mizumoto, K., Tanaka, M. and Kawakita, M. (1969). Purification and properties of factor G. Cold Spring Harbor Symp. Quant. Biol. *34*, 385-393.

Kirschbaum, J. B. and Scaife, S. (1974). Evidence for a transducing phage carrying the genes for β and β' subunits of *Escherichia coli* RNA polymerase. Mol. Gen. Genet. *132*, 193-201.

Laursen, R. A., Nagarkatti, S. and Miller, D. L. (1977). Amino acid sequence of elongation factor Tu. Characterization and alignment of the cyanogen bromide fragments and location of the cysteine residues. FEBS Letters *80*, 103-106.

Leder, P., Skogerson, L. E. and Nan, M. M. (1969). Translocation of mRNA codons. I. The preparation and characterization of a homogeneous enzyme. Proc. Nat. Acad. Sci. USA *62*, 454-460.

Lengyel, P. (1974). The process of translation. In Ribosomes, M. Nomura, A. Tissièrres and P. Lengyl, eds., Cold Spring Harbor Monograph Series, pp. 13-52.

Lindahl, L., Jaskunas, S. R., Dennis, P. P. and Nomura, M. (1975). Cluster of genes in *Escherichia coli* for ribosomal proteins, ribosomal RNA and RNA polymerase subunits. Proc. Nat. Acad. Sci. USA *72*, 2743-2747.

Lucas-Lenard, J. Cited in Lucas-Lenard, J. and Lippmann, F. (1971). Protein biosynthesis. Ann. Rev. Biochem. *40*, 409-448.

Nakamura, S., Arai, K., Takahashi, K. and Kaziro, Y. (1977). Alignment of the tryptic fragments and location of sulfhydryl groups of the polypeptide chain elongation factor Tu. Biochem. Biophys. Res. Commun. *77*, 1418-1424.

Neidhardt, F. C., Bloch, P. L., Pedersen, S. and Reeh, S. (1977). Chemical measurement of steady state levels of ten aminoacyl-transfer ribonucleic acid synthetases in *Escherichia coli*. J. Bacteriol. *129*, 378-387.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. *250*, 4007-4021.

Ovchinnikov, Y. A., Lipkin, V. M., Kodyanov, S., Chertov, O. and Smirnov, Y. V. (1977). Primary structure of α-subunit of DNA dependent RNA polymerase from *Escherichia coli*. FEBS Letters *76*, 103-112.

Pedersen, S., Reeh, S. R., Parker, J., Watson, R. J., Friesen, J. D. and Fiil, N. P. (1976). Analysis of the proteins synthesized in ultraviolet light-irradiated *Escherichia coli* following infection with the bacteriophage λdrif^{d}18 and λdfus-3. Mol. Gen. Genet. *144*, 339-344.

Subramanian, A. R., Haase, C. and Geisen, M. (1976). Isolation and characterization of a growth-cycle-reflecting, high-molecular-weight protein, associated with *Escherichia coli* ribosomes. Eur. J. Biochem. *67*, 591-601.

Wahba, A. J., Miller, M. J., Niveleau, A., Landers, T. A., Carmichael, G. G., Weber, K., Hawley, D. A. and Slobin, L. I. (1974). Subunit I of Qβ replicase and 30S ribosomal protein S1 of *Escherichia coli:* evidence for the identity of the two proteins. J. Biol. Chem. *249*, 3314-3316.

Watson, R. J., Parker, J., Fiil, N. P., Flaks, S. G. and Friesen, J. D. (1975). New chromosomal location for structural genes of ribosomal proteins. Proc. Nat. Acad. Sci. USA *72*, 2765-2769.

Wittinghofer, A. and Leberman, R. (1976). Elongation factor T from *Bacillus stearothermophilus* and *Escherichia coli*. Eur. J. Biochem. *62*, 373-382.

Wittmann, H. G. (1974). Purification and identification of *Escherichia coli* ribosomal proteins. In Ribosomes, M. Nomura, A. Tissièrres and P. Lengyl, eds., Cold Spring Harbor Monograph Series, pp. 93-114.

Wittmann-Liebold, B., Robinson, S. M. L. and Dzionara, M. (1977). Prediction of secondary structures in proteins from the *Escherichia coli* 30S ribosomal subunit. FEBS Letters *77*, 301-307.

Zabin, I. and Fowler, A. V. (1970). β-Galactosidase and thiogalactoside transformylase. In The Lactose Operon, J. R. Beckwith and D. Zipser, eds., Cold Spring Harbor Monograph Series, pp. 27-47.