The EMBO Journal vol.3 no.10 pp.2315–2318, 1984

Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells

Gunnar von Heijne

Research Group for Theoretical Biophysics, Department of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Communicated by K. Simons

A statistical analysis of the distribution of charged residues in the N-terminal region of 39 prokaryotic and 134 eukaryotic signal sequences reveals a remarkable similarity between the two samples, both in terms of net charge and in terms of the position of charged residues within the N-terminal region, and suggests that the formyl group on Met_f is not removed in prokaryotic signal sequences.

Key words: amino acid statistics/protein export/signal sequences

Introduction

Secreted proteins are generally synthesized with an N-terminal signal sequence, some 15–25 amino acids long, that somehow initiates the export process. Several functions seem to be encoded within this short peptide: in eukaryotes, it mediates the interaction between the ribosome and the so-called signal recognition particle (SRP) which arrests translation until an unoccupied export-site on the membrane is found (Walter *et al.*, 1981) [for a discussion of a possible prokaryotic SRP, see Kumamoto *et al.* (1984)]; it may influence the 'SRP-receptor'-catalyzed release of SRP from the ribosome (Gilmore and Blobel, 1983; Hall *et al.*, 1983; Silhavy *et al.*, 1983); and it contains the information needed for its removal from the mature protein.

These different functions most likely reside in at least partially non-overlapping regions of the signal sequence (Figure 1). A C-terminal region of about five or six residues seems to be involved in defining the site of cleavage between the signal sequence and the mature protein (von Heijne, 1983, 1984; Perlman and Halvorson, 1983); the hydrophobic core of the signal sequence has been implicated in binding to SRP (Walter *et al.*, 1981); and the net charge of the N-terminal region has been shown to influence both the level of translation and the efficiency of export (Hall *et al.*, 1983; Vlasuk *et al.*, 1983).

Although the functions performed by the signal sequence are understood in a general sense, it has proved difficult to correlate them with either the amino acid sequence or higher order structures. In part, this is because the primary sequence has been found to be remarkably variable, and a very large number of sequences have to be analyzed if any significant patterns of amino acids, beyond the overall tri-partite design alluded to above, are to be found. So far, only a few studies of this kind have been published (von Heijne, 1983, 1984; Perlman and Halvorson, 1983), dealing chiefly with the amino acid pattern around the cleavage site between the signal sequence and the mature protein.

Here, the statistical distribution of charged amino acids in the N-terminal region in a large sample of known signal sequences is analyzed for the first time. The results show that prokaryotic and eukaryotic signal sequences, although superficially not very similar in this region, nevertheless seem to be under comparable selective pressures, yet another instance of the apparently very close homology between the two export machineries. The observed sequence differences can possibly be attributed to the fact that prokaryotic proteins are made with a formylated and uncharged N-terminal methionine, whereas eukaryotic proteins are initiated by an unformylated, positively charged Met residue.

Results and Discussion

The number of signal sequences with a given N-terminal net charge (counting +1 for Arg and Lys, and −1 for Asp and Glu) is given in Table I. The distribution peaks at +1 for the eukaryotic sample and at +2 for the prokaryotes. Interestingly, the relative frequencies of N termini with a given net charge in the two samples overlap remarkably well if one positive charge is added to the eukaryotic sequences, Figure 2.

Likewise, the length distributions for the charged N-terminal regions have similar shapes (including only sequences with at least one charged residue besides the initiator Met in the samples, since the delineation of the polar N-terminal region is somewhat arbitrary in sequences lacking charged residues), except that the eukaryotic distribution peaks at a length of two residues, whereas the prokaryotic one has its main weight at a length of three residues, Figure 3.

An obvious explanation for these overlaps is that the selective constraints in terms of net charge on the prokaryotic and eukaryotic signal sequence N termini in fact are very similar, and that the initiator methionine (which is formylated and hence uncharged in prokaryotes, but unmodified and hence carrying a positively charged free amino group in eukaryotes) is not deformylated and remains bound to the N termini of the signal sequences. In the prokaryotes, an additional positively charged residue would thus be required, making these sequences on average one residue longer.

It is perhaps a little surprising that the uncharged Met_f in the prokaryotes must be compensated for by an additional Arg or Lys, since for cytoplasmic proteins deformylation and removal of the N-terminal Met are quite fast and take place shortly after the N terminus emerges from the ribosome (Pine, 1969; Housman *et al.*, 1972) – most likely, interactions with the export machinery (such as binding to SRP or membrane) competes with these cleavages, and protects Met_f from the deformylase.

If, in addition to the net charge, the positions of the charged residues within the N-terminal region are important, and if it is true that Met_f is not deformylated, we might expect to see a difference in the distribution of the charged residues between prokaryotes and eukaryotes since the latter automatically have a positive charge right at the N terminus. To test this, the number of sequences with either Arg or Lys next to the initiator Met was counted, only including N-terminal regions

⊕ MET ⊖ ASP TYR TYR ⊕ ARG ⊕ LYS TYR ALA ALA ILE PHE LEU VAL THR LEU SER VAL PHE LEU HIS VAL LEU HIS SER ↓ ALA ...

Fig. 1. A 'typical' signal sequence (human chorionic gonadotropin, α subunit) with the hydrophobic core in boldface and the cleavage site marked by an arrow. Note the charged residues in the N-terminal region, which has a net charge of +2.

Table I. Number of signal sequences with a given net charge in the N-terminal region (counting +1 for Arg and Lys, and −1 for Asp and Glu) in the eukaryotic and prokaryotic samples

	Net charge							
	−2	−1	0	+1	+2	+3	+4	+5
Eukaryotes	2	8	40	68	13	2	1	0
Prokaryotes	0	0	0	7	24	5	1	2

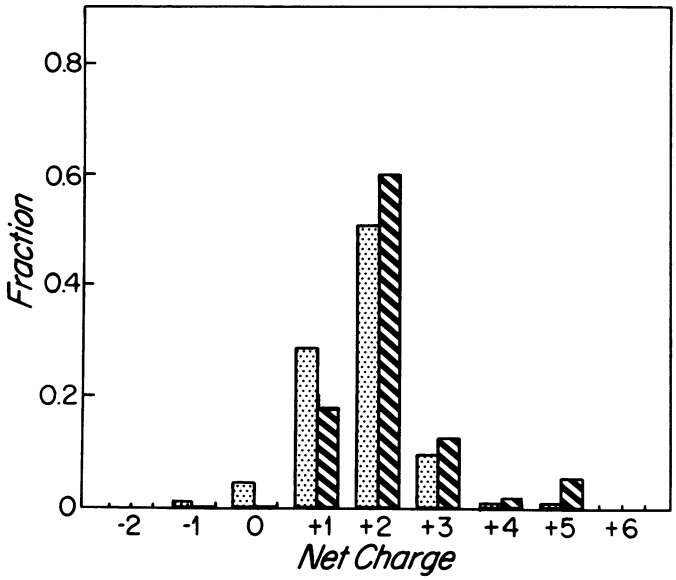

Fig. 2. Distribution of the net N-terminal charge calculated from Table I, with one positive charge added to each eukaryotic sequence. Eukaryotic sample = dotted bars, prokaryotic sample = hatched bars.

(defined as the segment from the initiator Met to the charged residue closest to the hydrophobic core) longer than three residues (eukaryotes) or four residues (prokaryotes) in the calculation. Amongst the eukaryotic sequences, only six out of 62 (10%) have Arg or Lys in this position, whereas there are 12 sequences out of 22 (55%) with a positively charged residue next to Met_f among the prokaryotic sequences. Thus, a positive charge close to the N terminus may be advantageous.

Similarly, in signal sequences with both positively and negatively charged residues in their N-terminal region (such as the one in Figure 1), the charged region ends with a positive charge on its C-terminal side in 13 out of 14 cases. This latter observation may be explained either as a functional requirement or simply as a consequence of selection against negatively charged residues: starting from a highly positively charged N-terminal region (the only kind that presumably will accept a negative charge), an acidic residue should be much more likely to be functionally neutral when placed well inside the charged region rather than in the hydrophobic region, C-terminally to the last positive charge, where the introduction of a charged residue may often have a deleterious effect

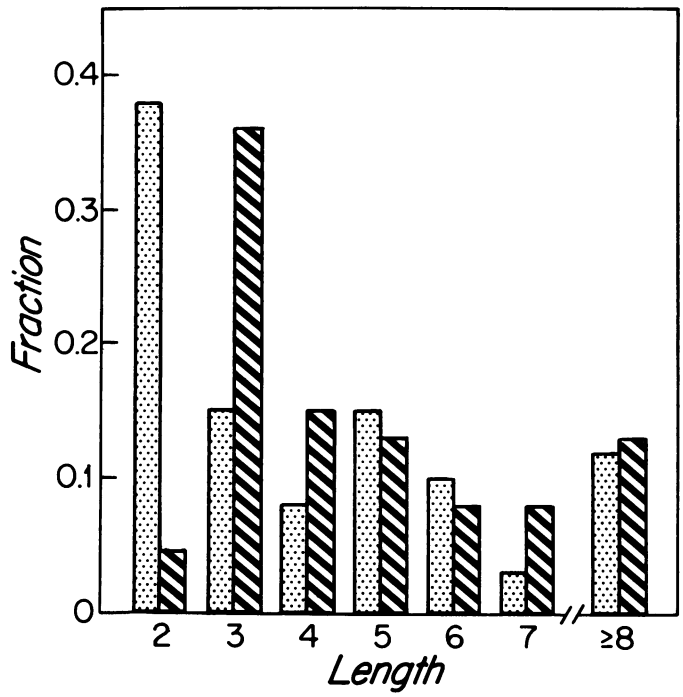

Fig. 3. Length distribution of the charged N-terminal regions (counting from the initiator Met to the charged residue closest to the hydrophobic core, see text). Eukaryotic sample = dotted bars, prokaryotic sample = hatched bars.

(Emr and Silhavy, 1982).

Although the connection between the selective constraints on the N-terminal region demonstrated here and the actual functioning of the signal sequence still remains to be made, one clue is provided by the work of Vlasuk *et al.* (1983) who have shown that both translation and secretion of the outer membrane lipoprotein from *Escherichia coli* are reduced when the net N-terminal charge is lowered from its wild-type value of +2: with zero net charge, synthesis is down to ~60% but secretion is unaffected; with a negative net charge, synthesis is reduced even further and, in addition, most of the protein made accumulates as precursor in the cytoplasm. This fits nicely with the net charge distribution in our sample, where 94% of all sequences have an N-terminal net charge of +1 or greater and 67% have a net charge of at least +2. There are only two sequences with a net charge less than zero, i.e., a κ-immunoglobulin light chain (L-321) (Burstein and Schechter, 1978), and pea seed lectin (Higgins *et al.*, 1983); unfortunately, not much is known about their level of synthesis and secretion.

Finally, if the explanation given above for the difference in the number of basic residues between the prokaryotic and eukaryotic signal sequences (i.e., the uncharged Met_f *versus* the positively charged Met at their respective N termini) is correct, one prediction is that archaebacterial signal sequences will turn out to be more like their eukaryotic than prokaryotic counterparts since they are synthesized with an unformylated initiator Met. In view of the results of Vlasuk *et*

al. referred to above, another implication is that up to 30% of all eukaryotic secreted proteins (those with a net charge of zero or less, not counting the initiator Met) may need one or more extra positively charged residues on their N terminus to be effectively made and exported from a prokyarotic cell.

Materials and methods

Thirty nine prokaryotic and 134 eukaryotic signal sequences are included in this study. Unless otherwise indicated, the original references can be found in Michaelis and Beckwith (1982) or von Heijne (1983, 1984).

Prokaryotic sequences: the sample includes all sequences listed in von Heijne (1984), plus the following: *Bacillus licheniformis* and *Staphylococcus aureus* penicillinase (Nielsen and Lampen, 1982); *E. coli* and *Erwinia amylovora* lipoprotein (Yamagata *et al.*, 1981); *E. coli* heat stable toxin ST1; *Corynebacterium diphtheriae* toxin (Kaczorek *et al.*, 1983); *E. coli* galactose-binding protein (Scripture and Hogg, 1983); phage IKe major and minor coat proteins (P.Peeters and R.Konings, personal communication); *Vibrio cholerae* toxin (Mekalanos *et al.*, 1983); *E. coli* papA, papC, papE, papF, papG, papH, pilin K88, and pilin K88a (M.Båga, S.Lundberg, B.Lund and S. Normark, personal communication); and molA (Clément and Hofnung, 1981).

Eukaryotic sequences: the sample includes all sequences listed in von Heijne (1984), except: mouse embryonic V_H immunoglobulin; rat $\varkappa$ immunoglobulin; human chorionic gonadotropin α subunit; mouse λ_L immunoglobulin; canine trypsinogen 2+3; human ϵ chain immunoglobulin; rabies virus glycoprotein CVS; and rat pituitary glycoprotein hormone α subunit. The sample also includes: murine and human α-fetoprotein (Law and Dugaiczyk, 1981; Morinaga *et al.*, 1983); human leukocyte interferon (Gray and Goeddel, 1982; Taniguchi *et al.*, 1980); human V_H, V_H-107, and H-chain (93g7) immunoglobulin (Bernstein *et al.*, 1982; Early *et al.*, 1980; Sims *et al.*, 1982); porcine gastrin; human influenza haemagglutinin A/Jap/ and avian influenza haemagglutinin A/FPV/; mouse $\varkappa_L$ immunoglobulins (L-41B, L-315, and L-321) (Burstein and Schechter, 1978); mouse embryonic V_H immunoglobulin; ovine α-S1 and α-S2 casein; mouse MHC H-$2L_d$ and E_β chains (Evans *et al.*, 1982; Saito *et al.*, 1983); mouse MHC I-A light chain (Malissen *et al.*, 1983); mouse MHC I_E-α-D chain (Hyldig-Nielsen *et al.*, 1983); mouse renin; catfish and anglerfish somatostatin (Magazin *et al.*, 1982; Argos *et al.*, 1983; Goodman *et al.*, 1980; Hobart *et al.*, 1980); human and porcine ACTH-β-LPH precursor; trypanosome surface glycoprotein; anglerfish glucagon; flounder antifreeze protein (Davies *et al.*, 1982); herpes simplex virus glycoproteins C and D (Frink *et al.*, 1983; Watson *et al.*, 1982); mouse C3 complement (Wiebauer *et al.*, 1982); bovine opiomelanocortin (Uhler and Herbert, 1983); human plasminogen activator (Pennica *et al.*, 1983); mouse epidermal growth factor (Gray *et al.*, 1983); pea seed lectin (Higgins *et al.*, 1983); rat apolipoprotein E (McLean *et al.*, 1983); rat prostatic binding protein C3 (Viskochil *et al.*, 1983); barley α-amylase (Rogers and Milliman, 1983); human apolipoprotein A-I (Karathanasis *et al.*, 1983); human HLA-DC α chain (Schenning *et al.*, 1984); human HLA-DC β chain (Larhammar *et al.*, 1983); human HLA-DS α chain (Chang *et al.*, 1983); human vasoactive intestinal polypeptide (Itoh *et al.*, 1983); mouse β-crystalline (Inana *et al.*, 1983); human interleukin II (Taniguchi *et al.*, 1983); bovine kininogen (Nawa *et al.*, 1983b); french bean phaseolin (Slightom *et al.*, 1983); rat angiotensinogen (Doolittle, 1983); yeast mating factor α-1, trypsinogen, rat elastase I and II, bovine chymosin, fruit fly vitellogenin I, human and bovine lactin (Dayhoff, 1983); bovine substance P (Nawa *et al.*, 1983a); mouse interleukin-3 (Fung *et al.*, 1984); *Amaranthus hybridus* psbA (Hirschenberg and McIntosh, 1983); pea legumin (Lycett *et al.*, 1984); *Aplysia* egg-laying hormone (Scheller *et al.*, 1983); human insulin-like growth factor (Jansen *et al.*, 1983); yeast K1 toxin (Skipper *et al.* 1984); Friend spleen focus-forming virus gp55 glycoprotein (Amanuma *et al.*, 1983); rat luteinizing hormone β-subunit (Chin *et al.*, 1983); rat seminal vesicle secretion IV protein (Harris *et al.*, 1983); human growth hormone-releasing factor (Gubler *et al.*, 1983); mouse T-cell receptor (Hedrick *et al.*, 1984); human T cell protein YT35 (Yanagi *et al.* 1984); rabbit poly-Ig receptor (Mostov *et al.*, 1984); and rat cholecystokinin (Deschenes *et al.*, 1984).

Acknowledgements

This work was carried out during a stay at the Department of Microbiology and Molecular Genetics, Harvard Medical School; my thanks to Professor Jon Beckwith for his support and hospitality, and to Dr. Kathy Strauch who taught me a lot about the biochemistry of Met_f.

References

Amanuma,H., Katori,A., Obata,M., Sagata,N. and Ikawa,Y. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3913-3917.

Argos,P., Taylor,W.L., Minth,C.D. and Dixon,J.E. (1983) *J. Biol. Chem.*, **258**, 8788-8793.

Bernstein,K.E., Premkumar Reddy,E., Alexander,C.B. and Mage,R.G. (1982) *Nature*, **300**, 74-76.

Burstein,Y. and Schechter,I. (1978) *Biochemistry (Wash.)*, **17**, 2392-2400.

Chang,H.-C., Moriuchi,T. and Silver,J. (1983) *Nature*, **305**, 813-815.

Chin,W.W., Godine,J.E., Klein,D.R., Chang,A.S., Tan,L.K. and Habener, J.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4649-4653.

Clément,J.M. and Hofnung,M. (1981) *Cell*, **27**, 507-514.

Davies,P.L., Roach,A.H. and Hew,C.-L. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 335-339.

Dayhoff,M. (1983) *Protein Sequence Data Bank.*

Deschenes,R.J., Lorentz,L.J., Haun,R.S., Roos,B.A., Collier,K.J. and Dixon,J.E. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 726-730.

Doolittle,R.F. (1983) *Science (Wash.)*, **222**, 417-419.

Early,P., Huang,H., Davis,M., Calame,K. and Hood,L. (1980) *Cell*, **19**, 981-992.

Emr,S.D. and Silhavy,T.J. (1982) *J. Cell Biol.*, **95**, 689-696.

Evans,G.A., Margulies,D.H., Camerini-Otero,R.D., Ozato,K. and Seidman, J.G. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1994-1998.

Frink,R.J., Eisenberg,R., Cohen,G. and Wagner,E.K. (1983) *J. Virol.*, **45**, 634-647.

Fung,M.C., Hapel,A.J., Ymer,S., Cohen,D.R., Johnson,R.M., Campbell, H.D. and Young,I.G. (1984) *Nature*, **307**, 233-237.

Furutani,Y., Morimoto,Y., Shibahara,S., Noda,M., Takahashi,H., Hirose, T., Asai,M., Inayama,S., Hayashida,H., Miyata,T. and Numa,S. (1983) *Nature*, **301**, 537-540.

Gilmore,R. and Blobel,G. (1983) *Cell*, **35**, 677-685.

Goodman,R.H., Jacobs,J.W., Chin,W.W., Lund,P.K., Dee,P.C. and Habener,J.F. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5869-5873.

Gray,A., Dull,T. and Ullrich,A. (1983) *Nature*, **303**, 722-725.

Gray,P.W. and Goeddel,D.V. (1982) *Nature*, **298**, 859-863.

Gubler,U., Monahan,J.J., Lomedico,P.T., Bhatt,R.S., Collier,K.J., Hoffman,B.J., Böhlen,P., Esch,F., Ling,N., Zeytin,F., Brazeau,P., Poonian, M.S. and Gage,L.P. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4311-4314.

Hall,M.N., Gabay,J. and Schwartz,M. (1983) *EMBO J.*, **2**, 15-19.

Harris,S.E., Mansson,P.-E., Tully,D.B. and Burkhart,B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6460-6464.

Hedrick,S.M., Nielsen,E.A., Kavaler,J., Cohen,D.I. and Davis,M.M. (1984) *Nature*, **308**, 153-158.

Higgins,T.J.V., Chandler,P.M., Zurawski,G., Button,S.C. and Spencer,D. (1983) *J. Biol. Chem.*, **258**, 9544-9549.

Hirschenberg,J. and McIntosh,L. (1983) *Science (Wash.)*, **222**, 1346-1349.

Hobart,P., Crawford,R., Shen,L., Pictet,R. and Rutter,W.J. (1980) *Nature*, **288**, 137-141.

Housman,D., Gillespie,D. and Lodish,H.F. (1972) *J. Mol. Biol.*, **65**, 163-166.

Hyldig-Nielsen,J.J., Schenning,L., Hammerling,U., Widmark,E., Heldin,E., Lind,P., Servenius,B., Lund,T., Flavell,R., Lee,J.S., Trowsdale,J., Schreier,P.H., Zablitsky,F., Larhammar,D., Peterson,P.A. and Rask.L. (1983) *Nucleic Acids Res.*, **11**, 5055-5071.

Inana,G., Piatigorsky,J., Norman,B., Slingsby,C. and Blundell,T. (1983) *Nature*, **302**, 310-315.

Itoh,N., Obata,K., Yanaihara,N. and Okamoto,H. (1983) *Nature*, **304**, 547-549.

Jansen,M., van Schaik,F.M.A., Ricker,A.T., Bullock,B., Woods,D.E., Gabbay,K.H., Nussbaum,A.L., Sussenbach,K.H. and Van der Brande, J.L. (1983) *Nature*, **306**, 609-611.

Kaczorek,M., Delpeyroux,F., Chenciner,N., Sreeck,R.E., Murphy,J.R., Boquet,P. and Tiollais,P. (1983) *Science (Wash.)*, **221**, 855-858.

Karathanasis,S.K., Zannis,V.I. and Breslow,J.L. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6147-6151.

Kumamoto,C.A., Oliver,D.B. and Beckwith,J. (1984) *Nature*, **308**, 863-864.

Larhammar,D., Hyldig-Nielsen,J.J., Servenius,B., Andersson,G., Rask.L. and Peterson,P.A. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7313-7317.

Law,S.W. and Dugaiczyk,A. (1981) *Nature*, **291**, 201-205.

Lycett,G.W., Croy,R.R.D., Shirsat,A.H. and Boulter,D. (1984) *Nucleic Acids Res.*, in press.

Magazin,M., Minth,C.D., Funckes,C.L., Deschenes,R., Tavianini,M.A. and Dixon,J.E. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5152-5156.

Malissen,M., Hunkapiller,T. and Hood,L. (1983) *Science (Wash.)*, **221**, 750-754.

McLean,J.W., Fukazawa,C. and Taylor,J.M. (1983) *J. Biol. Chem.*, **258**, 8993-9000.

Mekalanos,J.J., Swartz,D.J., Pearson,G.D.N., Harford,N., Groyne,F. and de Wilde,M. (1983) *Nature*, **306**, 551-557.

Michaelis,S. and Beckwith,J. (1982) *Annu. Rev. Microbiol.*, **36**, 435-465.

Morinaga,T., Sakai,M., Wegmann,T.G. and Tamaoki,T. (1983) *Proc. Natl.*

Acad. Sci. USA, **80**, 4604-4608.
Mostov,K.E., Friedlander,M. and Blobel,G. (1984) *Nature,* **308**, 37-43.
Nawa,H., Hirose,T., Takashima,H., Inayama,S. and Nakanishi,S. (1983a) *Nature,* **306**, 32-36.
Nawa,H., Kitamura,N., Hirose,T., Asai,M., Inayama,S. and Nakanishi,S. (1983b) *Proc. Natl. Acad. Sci. USA,* **80**, 90-94.
Nielsen,J.B.K. and Lampen,J.O. (1982) *J. Biol. Chem.,* **257**, 4490-4495.
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. and Goeddel,D.V. (1983) *Nature,* **301**, 214-221.
Perlman,D. and Halvorson,H.O. (1983) *J. Mol. Biol.,* **167**, 391-409.
Pine,M.J. (1969) *Biochim. Biophys. Acta,* **174**, 359-372.
Rogers,J.C. and Milliman,C. (1983) *J. Biol. Chem.,* **258**, 8169-8174.
Saito,H., Maki,R.A., Clayton,L.K. and Tonegawa,S. (1983) *Proc. Natl. Acad. Sci. USA,* **80**, 5520-5524.
Scheller,R.H., Jackson,J.F., McAllister,L.B., Rothman,B.S., Mayeri,E. and Axel,R. (1983) *Cell,* **32**, 7-22.
Schenning,L., Larhammar,D., Bill,P., Wiman,K., Jonsson,A.-K., Rask,L. and Peterson,P.A. (1984) *EMBO J.,* **3**, 447-452.
Scripture,J.B. and Hogg,R.W. (1983) *J. Biol. Chem.,* **258**, 10853-10855.
Silhavy,T.J., Benson,S.A. and Emr,S.D. (1983) *Microbiol. Rev.,* **47**, 313-344.
Sims,J., Rabbitts,T.H, Estess,P., Slaughter,C., Tucker,P.W. and Capra, J.D. (1982) *Science (Wash.),* **216**, 309-310.
Skipper,N., Thomas,D.Y. and Lau,P.C.K. (1984) *EMBO J.,* **3**, 107-111.
Slightom,J.L., Sun,S.M. and Hall,T.C. (1983) *Proc. Natl. Acad. Sci. USA,* **80**, 1897-1901.
Taniguchi,T., Mantei,N., Schwarzstein,M., Nagata,S., Muramatsu,M. and Weissmann,C. (1980) *Nature,* **285**, 547-549.
Taniguchi,T., Matsui,H., Fujita,T., Takaoka,C., Kashima,N., Yoshimoto, R. and Hamuro,J. (1983) *Nature,* **302**, 305-310.
Uhler,M. and Herbert,E. (1983) *J. Biol. Chem.,* **258**, 257-261.
Viskochil,D.H., Perry,S.T., Lea,O.A., Stafford,D.W., Wilson,E.M. and French,F.S. (1983) *J. Biol. Chem.,* **258**, 8861-8866.
Vlasuk,G.P., Inouye,S., Ito,H., Itakura,K. and Inouye,M. (1983) *J. Biol. Chem.,* **258**, 7141-7148.
von Heijne,G. (1983) *Eur. J. Biochem.,* **133**, 17-21.
von Heijne,G. (1984) *J. Mol. Biol.,* **173**, 243-251.
Walter,P., Ibrahimi,I. and Blobel,G. (1981) *J. Cell Biol.,* **91**, 545-550.
Watson,R.J., Weis,J.H., Salstrom,J.S. and Enquist,L.W. (1982) *Science (Wash.),* **218**, 381-384.
Wiebauer,K., Domdey,H., Diggelmann,H. and Fey,G. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 7077-7081.
Yamagata,H., Nakamura,K. and Inouye,M. (1981) *J. Biol. Chem.,* **256**, 2194-2198.
Yanagi,Y., Yoshikai,Y., Leggett,K., Clark,S.P., Aleksander,I. and Mak, T.W. (1984) *Nature,* **308**, 145-149.

Received on 29 June 1984