Volume 303, number 2,3, 141–146 FEBS 11077 June 1992


# The amino acid composition is different between the cytoplasmic and extracellular sides in membrane proteins

Hiroshi Nakashima[a] and Ken Nishikawa[b]

[a]*School of Allied Medical Professions, Kanazawa University, 5-11-80 Kodatsuno, Kanazawa 920, Japan* and [b]*Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan*

Received 19 September 1991; revised version received 31 March 1992

The amino acid composition of transmembrane proteins was analyzed for their three separate portions: the transmembrane apolar, cytoplasmic and extracellular regions. The composition was different between cytoplasmic and extracellular peptides; alanine and arginine residues were preferentially sited on the cytoplasmic side, while the threonine and cysteine/cystine were preferentially sited on the extracellular side. The composition of cytoplasmic and extracellular peptides of membrane proteins corresponded to those of intracellular and extracellular types of soluble proteins, respectively. This difference in composition was independent of the peptide orientation against the membrane. Peptide chains could be correctly assigned as either cytoplasmic or extracellular, solely from an analysis of sequence composition. For single-spanning membrane proteins the predictive accuracy was 90%, whereas for multi-spanning proteins this was 85%.

Amino acid composition; Membrane protein; Transmembrane topology; Prediction

## 1. INTRODUCTION

The analysis of an amino acid sequence is a simple and useful way to classify proteins into groups. Dayhoff [1] originally classified proteins into families or superfamilies using standard alignment procedures and, more recently, about 8,000 proteins from the Protein Identification Resource (PIR) database [2] were classified into around 2,500 superfamilies. We have reported [3–6] that the character of a protein is also related to its amino acid composition and this varies with location (inside or outside of the cell), function (enzyme or non-enzyme) and folding type. The amino acid composition of an extracellular protein, which is characterized as having a signal peptide at the N-terminus, is different from that of intracellular proteins [4].

Here we have analyzed membrane proteins to see whether a difference in amino acid composition also occurs between cytoplasmic and extracellular portions. We initially focused our analysis on single-spanning membrane proteins because their chain constitution is simple. Furthermore, the location and orientation of the single transmembrane region can be determined with greater accuracy be experimental means than multi-spanning ones. The amino acid composition was analyzed for three distinct regions: the cytoplasmic and extracellular domains, as well as the transmembrane segment. Those proteins which have reverse orientations to the membrane were also included in order to examine the effect of peptide chain directionality. A similar analysis was then applied to multi-spanning membrane proteins to ascertain whether the trends observed in single-spanning membrane proteins held true for all membrane proteins.

*Correspondence address:* H. Nakashima, Protein Engineering Research Institute, 6-2-3 Furuedai, Suita, Osaka 565, Japan. Fax: (81) (762) 61-5281.

## 2. MATERIALS AND METHODS

### 2.1. *Protein data*

The 45 single-spanning and 24 multi-spanning membrane proteins were taken from release 27.0 of the PIR database [2]. The first 30 single-spanning proteins listed in Table I have been classified into superfamilies [2], while the remainder have yet to be classified. Most are plasma membrane proteins such as hormone receptors or antigen proteins. Others are endoplasmic reticulum membrane proteins comprising 38 from eukaryotes, 4 from prokaryotes and 3 from viruses. Only proteins with less than 25% residue identity were included in this analysis. Information on transmembrane location and orientation was obtained from the annotations of the PIR database or the literature cited in Table I.

The classification of transmembrane proteins proposed by von Heijne and Gavel [7] was adopted. A type-I protein has a signal peptide for secretion, which is cleaved upon maturation. The N-terminus of the mature protein sits on the extracellular side, whereas its C-terminus is in the cytoplasm. A type-II protein has the opposite orientation to type I and no signal peptide. A type-III protein has the same orientation as type-I but has no cleavable signal peptide. In this work 32 type-I, 8 type-II and 5 type-III proteins were used as examples of single-spanning membrane proteins (Table I).

Multi-spanning transmembrane proteins are classified as type-IV [7]. 24 examples are shown in the footnote of Table III together with their PIR codes. They comprise 18 eukaryotic, 4 prokaryotic and 2 virus proteins.

### 2.2. *Calculation of composition and distance*

Since the amino acid composition of short peptides can vary signif-

icantly we have examined only those cytoplasmic and extracellular peptides longer than 50 residues, excluding the cleavable signal peptide. In total, 73 peptides in single-spanning proteins and 54 peptides in multi-spanning proteins were used. The number of residues in each cytoplasmic and extracellular domain examined is shown in the columns of $N_{cyt}$ and $N_{ext}$ in Table I.

A protein or peptide can be represented as one point in amino acid composition space [3–6]. Two proteins (or peptides) are compared by measuring the distance between them. The normalized composition is defined as $CA_i = A_i/SD_i$, where $A_i$ is the content (in percentage) of amino acid type-i and $SD_i$ is the standard deviation over a large set of proteins. The distance between two peptides, j and k, is thus given by:

$$d_{jk} = [\sum_{i=1}^{20} (CA_{ij} - CA_{ik})^2]^{\frac{1}{2}}$$

The average composition ($AV_i$) and standard deviation ($SD_i$) listed in Table II were calculated using a large number of proteins [6].

## 3. RESULTS AND DISCUSSION

The average amino acid composition of the cytoplas-

Table I

45 single-spanning transmembrane proteins listed with their PIR codes

| No. | Code | Protein | Type | $N_{cyt}$ | $N_{ext}$ | Ref. |
|---|---|---|---|---|---|---|
| 1 | URXLA2 | Alpha-amidating enzyme II – African clawed frog | I | 88 | 724 | 8 |
| 2 | RDRT04 | NADPH cytochrome P-450 reductase – Rat | III | 634 | | 9, 10 |
| 3 | *GQHUE | Epidermal growth factor receptor – Human | I | 542** | 621 | |
| 4 | NMIV | Sialidase – Influenza A virus | II | | 419 | 11 |
| 5 | LWECB | $H^+$-transporting ATP synthase b chain – *Escherichia coli* | III | 130 | | 12 |
| 6 | PWRYNB | $Na^+$, $K^+$-transporting ATPase beta chain – Electric ray | II | | 242 | 13 |
| 7 | OYURGA | Guanylate cyclase – Sea urchin | I | 458 | 478 | |
| 8 | *QRHULD | LDL receptor – Human | I | 50** | 767 | |
| 9 | *GQHUN | Nerve growth factor receptor – Human | I | 155 | 222 | |
| 10 | UHHU2 | Interleukin-2 receptor – Human | I | | 219 | |
| 11 | *QRRBG | Poly-Ig receptor – Rabbit | I | 103 | 629 | |
| 12 | *BNRT3 | Myelin-associated glycoprotein, long form – Rat | I | 90 | 494 | |
| 13 | TDRTOX | OX-2 membrane glycoprotein – Rat | I | | 202 | |
| 14 | RWHUC2 | T-cell surface glycoprotein CD2 – Human | I | 115 | 182 | |
| 15 | HLHUDA | Class II histocompatibility antigen alpha chain – Human | I | | 191 | |
| 16 | HLHUG | Class II histocompatibility antigen gamma chain – Human | II | | 160** | |
| 17 | RWHUD1 | T-cell surface glycoprotein CD3 delta chain – Human | I | | 79 | |
| 18 | *TDRTLT | Leukocyte common antigen – Rat (fragment) | I | 705 | | |
| 19 | SAHU4F | Cell surface antigen 4F2 heavy chain – Human | II | 81 | 425** | |
| 20 | QRHUA4 | Alzheimer's disease amyloid A4 protein – Human | I | | 607** | |
| 21 | MPRT0 | Myelin P0 protein – Rat | I | 69 | 124 | |
| 22 | LNHU2A | Hepatic lectin H2a – Human | II | 58 | 233 | |
| 23 | GFHUE | Glycophorin A – Human | I | | 72 | |
| 24 | *NBHUIA | Platelet glycoprotein Ib alpha chain – Human | I | 96 | 485 | |
| 25 | IJHULM | Leukocyte adhesion protein beta chain – Human | I | | 678 | |
| 26 | JXHU | Transferrin receptor – Human | II | 61 | 671 | |
| 27 | *TFHUM | Melanotransferrin – Human | I | | 694 | |
| 28 | QRECS | Serine chemoreceptor protein – *Escherichia coli* | I | 323 | 151** | 14 |
| 29 | *HNVZVV | Hemagglutinin – Vaccinia virus | I | | 259 | |
| 30 | MMIV2 | Matrix (M2) protein – Influenza A virus | III | 54 | | 15 |
| 31 | A25970 | Transcriptional activator protein – *Vibrio cholerae* | II | 182 | 96 | 16 |
| 32 | A22841 | Reaction center H chain – *Rhodopseudomonas viridis* | III | 223 | | 17 |
| 33 | *S00670 | Gene Delta protein – Fruit fly | I | 261 | 572 | 18 |
| 34 | S03900 | Gene torso protein – Fruit fly | I | 503 | 379 | 19 |
| 35 | A31932 | IgG Fc receptor – Human | I | 76 | 178 | 20 |
| 36 | A30342 | Interleukin-2 receptor beta chain – Human | I | 286 | 214 | 21 |
| 37 | A26168 | Ribophorin I – Human | I | 150 | 416 | 22 |
| 38 | B26168 | Ribophorin II – Human | I | 70** | 517 | 22 |
| 39 | S04530 | Somatotropin receptor, hepatic – Human | I | 350 | 246 | 23 |
| 40 | *A26396 | T-cell surface glycoprotein CD5 – Human | I | 93 | 347 | 24 |
| 41 | A27068 | Cation-dependent mannose 6-phosphate receptor – Bovine | I | 67 | 165 | 25 |
| 42 | *A30788 | Cation-independent mannose 6-phosphate receptor – Bovine | I | 163 | 2269 | 26 |
| 43 | *S05479 | Neural cell adhesion molecule – Mouse | I | 114 | 1104 | 27 |
| 44 | A27410 | Plasma cell membrane protein – Mouse | II | 58 | 826 | 28 |
| 45 | A25742 | Platelet-derived growth factor receptor – Mouse | I | 542 | 499 | 29 |

'Type' indicates the type of membrane protein as defined by von Heijne and Gavel [7]. '$N_{cyt}$' and '$N_{ext}$' indicate the number of residues of a peptide on the cytoplasmic and extracellular sides, respectively. Those peptides of less than 50 residues long which were not used in the analysis are left blank. Proteins marked with * before the code contain duplication in their sequences. Peptides marked with ** are those which failed the assignment based on the composition data.

mic (CYT), extracellular (EXT) and apolar membrane (MEM) regions of single-spanning membrane proteins are shown in Table II. These were calculated using 33, 40 and 45 proteins, respectively (Table I). Compared to the average (AV) composition calculated from proteins of the PIR database [6], the content of hydrophobic residues was low and that of hydrophilic residues was high for both cytoplasmic and extracellular regions. This is because the average calculated from the PIR database includes a considerable number of proteins with transmembrane regions, whereas our cytoplasmic and extracellular peptides are without them. The transmembrane region was rich in hydrophobic residues such as Leu, Ile, and Val (the three letter nomenclature for amino acids is used hereafter), and this was similar to other analyses [6,7,30,31].

Table II shows deviation in the amino acid composition between the CYT and EXT regions. It is interesting to note that $\alpha$-helix-favoring residues [32,33] such as Glu, Met, Ala and Leu are more frequently observed in the peptide on the cytoplasmic side, while $\beta$-sheet-favoring residues [32,33], such as Val, Ile, Phe and Tyr, are more frequently found on the extracellular side. Charged residues, both basic (Arg, Lys) and acidic (Glu, Asp) are preferentially sited on the cytoplasmic side whereas aromatic residues (Trp, Tyr, Phe) are preferentially found on the extracellular side. For histidine, which is aromatic in terms of character and charge, there is no preference.

Table II

Average amino acid compositions (%) of cytoplasmic (CYT), extracellular (EXT) and membrane (MEM) regions of single-spanning transmembrane proteins

| Amino acid | CYT | (CYT2) | EXT | (EXT2) | MEM | AV | SD |
|---|---|---|---|---|---|---|---|
| Ala | 8.63 | (10.88) | 5.15 | (5.04) | 9.90 | 8.0 | 3.7 |
| Cys | 1.03 | (0.69) | 3.24 | (2.20) | 2.55 | 1.8 | 2.3 |
| Asp | 6.24 | (6.13) | 5.75 | (5.26) | 0.35 | 5.1 | 2.2 |
| Glu | 7.82 | (9.34) | 7.05 | (6.07) | 0.08 | 6.1 | 3.0 |
| Phe | 2.73 | (2.93) | 3.52 | (3.72) | 6.47 | 3.8 | 1.9 |
| Gly | 6.80 | (7.72) | 6.38 | (7.09) | 8.14 | 6.9 | 3.4 |
| His | 2.70 | (2.15) | 2.69 | (2.99) | 0.20 | 2.2 | 1.6 |
| Ile | 3.48 | (1.80) | 4.40 | (4.32) | 15.25 | 5.5 | 2.5 |
| Lys | 6.25 | (6.11) | 5.25 | (6.31) | 0.16 | 6.0 | 3.4 |
| Leu | 8.44 | (8.03) | 8.11 | (9.88) | 22.28 | 9.2 | 3.4 |
| Met | 2.14 | (3.79) | 1.60 | (1.85) | 1.85 | 2.5 | 1.4 |
| Asn | 4.18 | (5.75) | 4.81 | (5.94) | 0.94 | 4.3 | 2.3 |
| Pro | 6.28 | (7.21) | 5.65 | (6.22) | 2.38 | 5.0 | 3.2 |
| Gln | 4.76 | (4.68) | 4.45 | (4.50) | 0.87 | 4.0 | 2.4 |
| Arg | 6.75 | (6.01) | 4.38 | (3.73) | 0.09 | 5.9 | 3.3 |
| Ser | 8.53 | (7.25) | 8.04 | (8.05) | 4.17 | 6.8 | 2.8 |
| Thr | 4.43 | (3.51) | 7.41 | (5.20) | 4.33 | 5.8 | 2.6 |
| Val | 5.44 | (4.57) | 7.00 | (6.19) | 14.34 | 6.7 | 2.5 |
| Trp | 0.80 | (0.47) | 1.68 | (2.10) | 2.21 | 1.3 | 1.2 |
| Tyr | 2.54 | (1.01) | 3.42 | (3.32) | 3.42 | 3.2 | 1.8 |
| No. of peptides | 33 | 5 | 40 | 8 | 45 | | |

CYT2 and EXT2 indicate composition of type-II proteins. The total number of peptides used in the estimation is indicated in the bottom row. Amino acid composition was estimated for individual peptides and then the average was calculated over all the peptides considered. AV and SD stand for average and standard deviation obtained from a large number of proteins [6].

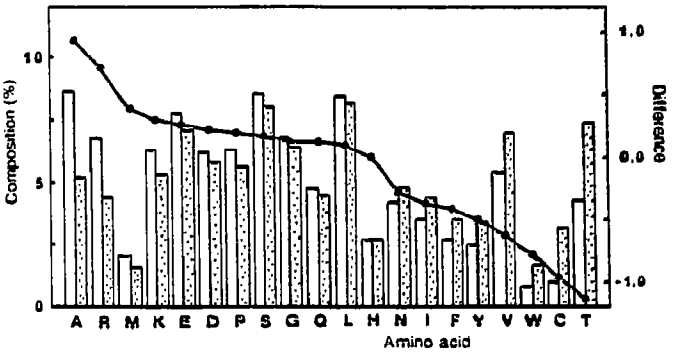


Fig. 1. Average amino acid composition of the cytoplasmic (in open bars) and extracellular domains (in dotted bars) of single-spanning transmembrane proteins. The difference calculated as (cytoplasmic-extracellular)/(standard deviation) is plotted with filled circles.

In order to show the deviation more clearly we calculated the difference defined as (CYT–EXT)/SD. The composition of both cytoplasmic and extracellular peptides are plotted in Fig. 1 in decreasing order of difference (CYT–EXT). As shown in Fig. 1, the siting of Ala and Arg residues shows a strong preference for the cytoplasmic side, while that for Thr, Cys and Trp residues shows a preference for the extracellular side. A similar deviation in amino acid content was observed between intracellular and extracellular types of soluble proteins [4].

The compositions of type-II membrane proteins, which have the reverse orientation, are shown separately in columns CYT2 and EXT2 of Table II. The average composition of type-II proteins does not differ significantly from those of cytoplasmic and extracellular peptides, indicating that the difference in the amino acid composition is independent of peptide orientation.

Next the deviation of amino acid content was examined for 28 single-spanning proteins that have peptides of more than 50 residues on both sides of the membrane (see Table I). For 22 of these proteins, Ala was more frequently found on the cytoplasmic side. Conversely, in 25 of the 28 proteins, Thr was found preferentially on the extracellular side. The deviation of amino acids was correlated with the average composition of peptide chains on the cytoplasmic and extracellular sides.

Each peptide in Table I (33 cytoplasmic and 40 extracellular) was characterized using the distance in composition space. Distances for a given peptide to the points defined by the average compositions of CYT and EXT (Table II) were used to classify each peptide depending on which had the shorter distance. For instance, in the first protein in Table I, URXLA2, the

cytoplasmic peptide (88 residues) has distances of 5.53 and 6.09 to the average points of CYT and EXT, respectively. Thus, it is classified as CYT type because of the shorter distance to CYT. In other words, the peptide has a composition typical of a cytoplasmic peptide. The extracellular peptide (724 residues) of the same protein is also typical, having a closer distance to EXT. The set of distances in this case are CYT=2.88 and EXT=2.27. In this way, all the peptides of single-spanning proteins were examined and the results showed that 90% (66/73) of them have typical compositions. In other words, they can be correctly assigned as either of CYT or EXT types from their amino acid composition. Peptides incorrectly assigned are marked in Table I. It should be noted that the success was similar for type-II proteins where 11 peptides out of 13 (85%) were correctly assigned. This reinforces the previous statement and suggests that the characteristic composition of each domain is independent of the transmembrane topology.

The analysis was repeated for 24 multi-spanning membrane proteins. Average compositions for 34 cytoplasmic and 20 extracellular peptides (longer than 50 residues) are shown in Table III. The average composition of multi-spanning proteins does not differ significantly from those of single-spanning proteins (Table II). The assignment with distance in composition space was also applied to show that distinction between cytoplasmic or extracellular sides could successfully be made. For multi-spanning proteins an accuracy of 85% (46/54) was observed. These data indicate that the trend found in single-spanning proteins generally holds for membrane proteins, including those with long connecting loops between transmembrane segments.

Since the three-dimensional structures for proteins in Table I are not available, except for the photosynthetic reaction center H chain [17], the locations of transmembrane regions are not certain. Consequently we followed the authors' assignments given in the literature. In the present study, however, we are concerned with the amino acid composition of cytoplasmic and extracellular peptides composed of at least 50 residues, so deficiencies in the assignment of transmembrane regions should not affect the results significantly. The number of membrane proteins available from prokaryotes and viruses are not large enough to draw any firm conclusions. Some of the proteins in Table I contain duplicated sequences of longer than 50 residues, such as the LDL receptor which has repeated sequences of the EGF-like module [35]. We checked whether the amino acid composition might be biased towards duplication, and found that the average composition for 13 proteins with duplications did not differ significantly from the others in either the cytoplasmic or extracellular peptides.

It is known that carbohydrate attachment by *N*-glycosidic linkage occurs at Asn-X-Ser or Asn-X-Thr sites on the extracellular side. Since 82% (37/45) of single-spanning proteins used here were glycoproteins, the content of Asn, Ser and Thr residues had been expected to be higher in the extracellular side than in the cytoplasm. However, on the extracellular side the content of Asn was not as high as that of Thr, and Ser was found

Table III

Average amino acid compositions (%) of cytoplasmic (CYT), extracellular (EXT) and membrane (MEM) regions for 24 multi-spanning transmembrane proteins.

| Amino acid | CYT | EXT | MEM |
|---|---|---|---|
| Ala | 6.69 | 5.08 | 9.36 |
| Cys | 1.70 | 2.95 | 2.56 |
| Asp | 4.97 | 5.96 | 0.94 |
| Glu | 7.76 | 6.04 | 0.94 |
| Phe | 3.59 | 4.36 | 10.99 |
| Gly | 6.23 | 8.20 | 6.17 |
| His | 2.11 | 2.10 | 0.47 |
| Ile | 4.51 | 4.95 | 13.73 |
| Lys | 8.36 | 4.93 | 0.58 |
| Leu | 8.23 | 8.03 | 16.64 |
| Met | 2.46 | 2.61 | 3.93 |
| Asn | 4.49 | 5.75 | 2.31 |
| Pro | 5.20 | 4.84 | 1.96 |
| Gln | 5.39 | 4.24 | 1.14 |
| Arg | 6.65 | 4.75 | 0.27 |
| Ser | 7.40 | 6.41 | 5.58 |
| Thr | 5.18 | 5.87 | 4.68 |
| Val | 5.27 | 6.07 | 12.43 |
| Trp | 1.06 | 2.31 | 2.20 |
| Tyr | 2.75 | 4.55 | 3.13 |
| No. of peptides | 34 | 20 | 127 |

The total number of peptides used in the estimation is indicated at the bottom row. The 24 multi-spanning membrane proteins used in this analysis are listed below with protein name, PIR code, and regions of cytoplasmic or extracellular peptides: a region indicated in reverse order, e.g. 314–221 in the first protein below, represents an extracellular peptide from residues 221 to 314, and one in the normal order as 340–887 represents a cytoplasmic peptide from residues 340 to 887. Peptide regions shorter than 50 residues are not shown. Hydroxymethylglutaryl-CoA reductase (RDHYE 314–221, 340–887) [39]; $Na^+/K^+$-transporting ATPase α chain (PWSHNA 1–93, 148–288, 347–784 [40]; myelin proteolipid protein (MPBOPL 150–97) [41]; β-2-adrenergic receptor (QRHYB2 221–274, 330–418) [42]; nicotinic acetylcholine receptor α chain (ACHUA1 230–21, 319–428) [43]; sodium channel protein (CHEE 1–113, 378–265, 403–557, 766–712, 1235–1153, 1543–1475, 1568–1820) [44]; multidrug resistance protein 1 (DVHU1 1–51, 237–296, 778–832, 875–936) [45]; inner membrane protein *mal*F (MMECMF 276–92) [46]; virulence membrane protein *pho*Q (VZEBPT 219–487) [47]; *env* polyprotein (VCLJSP 1–63, 578–88) [48]; E1 membrane glycoprotein (MMIHIB 102–225) [49]; cytochrome *c* oxidase polypeptide II (S00106 297–130) [50]; reaction center protein M chain (B25102 2–52) [17]; membrane protein patched (S06119 1–73, 426–93, 586–676, 966–700) [51]; potassium channel protein A (S00479 1–227, 479–616) [52]; follicle stimulating hormone receptor (JN0122 631–695) [53]; glucose-transport protein (A27217 207–271) [54]; muscarinic acetylcholine receptor M2 (A27386 208–387) [55]; γ-aminobutyric acid/benzodiazepine receptor (B27142 245–26, 328–451) [56]; peripherin (A34608 263–122, 283–345) [57]; dihydrophyridine receptor (A30063 1–51, 309–219, 335–432, 636–582, 662–799, 1066–1118, 1356–1290) [58]; glutamate receptor (S07059 480–19, 633–805, 907–827) [59]; glycine receptor (A27141 315–398) [60]; and steroid 17-α-monooxygenase (A30828 22–168, 507–187) [61,62].

more often on the cytoplasmic side. Therefore, the high content of Thr on the extracellular side cannot be explained by glycosylation alone. The high content of Cys residues on the extracellular side may result from disulfide bond formation. Basic residues (Arg, Lys) were frequently observed at the edge of the cytoplasmic side as stop-transfer signals [7,36–39]. This may help to explain why Arg and Lys are more abundant in cytoplasmic peptides.

The reason why such an imbalance of the amino acid composition exists between the cytoplasmic and extracellular peptides is currently unexplained. One possible reason is that some kind of restriction in the sequence is required for transportation of a peptide chain through the lipid bilayer of the membrane. Nevertheless, this study shows that it is possible to infer for membrane proteins which peptide portion flanking a transmembrane segment is located on which side of the membrane.

## REFERENCES


[1] Dayhoff, M.O. (1976) Atlas of Protein Sequence and Structure, vol. 5, Suppl. 2, Natl. Biomed. Res. Foundation Silver Spring, Washington DC.
[2] Baker, W.C., George, D.G. and Hunt, L.T. (1990) Methods Enzymol. 183, 31–49.
[3] Nishikawa, K. and Ooi, T. (1982) J. Biochem. 91, 1821–1824.
[4] Nishikawa, K., Kubota, Y. and Ooi, T. (1983) J. Biochem. 94, 997–1007.
[5] Nakashima, H., Nishikawa, K. and Ooi, T. (1986) J. Biochem. 99, 153–162.
[6] Nakashima, H., Nishikawa, K. and Ooi, T. (1990) Proteins 8, 173–178.
[7] von Heijne, G. and Gavel, Y. (1988) Eur. J. Biochem. 174, 671–678.
[8] Ohsuye, K., Kitano, K., Wada, Y., Fuchimura, K., Tanaka, S., Mizuno, K. and Matsuo, H. (1988) Biochem. Biophys. Res. Commun. 150, 1275–1281.
[9] Black, S.D. and Coon, M.J. (1982) J. Biol. Chem. 257, 5929–5938.
[10] Porter, T.D. and Kasper, C.B. (1985) Proc. Natl. Acad. Sci. USA 82, 973–977.
[11] Fields, S., Winter, G. and Brownlee, G.G. (1981) Nature 290, 213–217.
[12] Hoppe, J. and Sebald, W. (1984) Biochem. Biophys. Acta 768, 1–27.
[13] Noguchi, S., Noda, M., Takahashi, H., Kawakami, K., Ohta, T., Nagano, K., Hirose, T., Inayama, S., Kawamura, M. and Numa, S. (1986) FEBS Lett. 196, 315–320.
[14] Boyd, A., Kendall, K. and Simon, M.I. (1983) Nature 301, 623–626.
[15] Lamb, R.A., Zebedee, S.L. and Richardson, C.D. (1985) Cell 40, 627–633.
[16] Miller, V.L., Taylor, R.K. and Mekalanos, J.J. (1987) Cell 48, 271–279.
[17] Deisenhofer, J., Epp, O., Miki, K., Huber, R. and Michel, H. (1985) Nature 318, 618–624.
[18] Vässin, H., Bremer, K.A., Knust, E. and Campos-Ortega, J.A. (1987) EMBO J. 6, 3431–3440.
[19] Sprenger, F., Stevens, L.M. and Nüsslein-Volhard, C. (1989) Nature 338, 478–483.
[20] Hibbs, M.L., Bonadonna, L., Scott, B.M., McKenzie, I.F.C. and Hogarth, P.M. (1988) Proc. Natl. Acad. Sci. USA 85, 2240–2244.
[21] Hatakeyama, M., Tsudo, M., Minamoto, S., Kono, T., Doi, T., Miyata, T., Miyasaka, M. and Taniguchi, T. (1989) Science 244, 551–556.
[22] Crimaudo, C., Hortsch, M., Gausepohl, H. and Meyer, D.I. (1987) EMBO J. 6, 75–82.
[23] Leung, D.W., Spencer, S.A., Cachianes, G., Hammonds, R.G., Collins, C., Henzel, W.J., Barnard, R., Waters, M.J. and Wood, W.I. (1987) Nature 330, 537–543.
[24] Jones, N.H., Clabby, M.L., Dialynas, D.P., Huang, H.-J.S., Herzenberg, L.A. and Strominger, J.L. (1986) Nature 323, 346–349.
[25] Dahms, N.M., Lobel, P., Breitmeyer, J., Chirgwin, J.M. and Kornfeld, S. (1987) Cell 50, 181–192.
[26] Lobel, P., Dahms, N.M. and Kornfeld, S. (1988) J. Biol. Chem. 263, 2563–2570.
[27] Moos, M., Tacke, R., Scherer, H., Teplow, D., Früh, K. and Schachner, M. (1988) Nature 334, 701–703.
[28] van Driel, I.R. and Goding, J.W. (1987) J. Biol. Chem. 262, 4882–4887.
[29] Yarden, Y., Escobedo, J.A., Kuang, W.-J., Yang-Feng, T.L., Daniel, T.O., Tremble, P.M., Chen, E.Y., Ando, M.E., Harkins, R.N., Francke, U., Fried, V.A., Ullrich, A. and Williams, L.T. (1986) Nature 323, 226–232.
[30] Kuhn, L.A. and Leigh Jr., J.S. (1985) Biochim. Biophys. Acta 828, 351–361.
[31] Deber, C.M., Brandl, C.J., Deber, R.B., Hsu, L.C. and Young, X.K. (1986) Arch. Biochem. Biophys. 251, 68–76.
[32] Chou, P.Y. and Fasman, G.D. (1978) Annu. Rev. Biochem. 47, 251–276.
[33] Levitt, M. (1978) Biochemistry 17, 4277–4285.
[34] Baron, M., Norman, D.G. and Campbell, I.D. (1991) Trends Biochem. Sci. 16, 13–17.
[35] Sabatini, D.D., Kreibich, G., Morimoto, T. and Adesnik, M. (1982) J. Cell Biol. 92, 1–22.
[36] von Heijne, G. (1986) EMBO J. 5, 3021–3027.
[37] von Heijne, G. (1989) Nature 341, 456–458.
[38] Hartmann, E., Rapoport, T.A. and Lodish, H.F. (1989) Proc. Natl. Acad. Sci. USA 86, 5786–5790.
[39] Liscum, L., Finer-Moore, J., Stroud, R.M., Luskey, K.L., Brown, M.S. and Goldstein, J.L. (1985) J. Biol. Chem. 260, 522–530.
[40] Ovchinnikov, Yu.A., Arzamazova, N.M., Arystarkhova, E.A., Gevondyan, N.M., Aldanova, N.A. and Modyanov, N.N. (1987) FEBS Lett. 217, 269–274.
[41] Stoffel, W., Hillen, H. and Giersiefen, H. (1984) Proc. Natl. Acad. Sci. USA 81, 5012–5016.
[42] Dohlman, H.G., Bouvier, M., Benovic, J.L., Caron, M.G. and Lefkowitz, R.J. (1987) J. Biol. Chem. 262, 14282–14288.
[43] Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Tanabe, T., Shimizu, S., Kikyotani, S., Kayano, T., Hirose, T., Inayama, S. and Numa, S. (1983) Nature 305, 818–823.
[44] Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T., Ikeda, T., Takahashi, H., Nakayama, H., Kanaoka, Y., Minamino, N., Kangawa, K., Matsuo, H., Raftery, M.A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. and Numa, S. (1984) Nature 312, 121–127.
[45] Chen, C., Chin, J.E., Ueda, K., Clark, D.P., Pastan, I., Gottesman, M.M. and Roninson, I.B. (1986) Cell 47, 381–389.
[46] Boyd, D., Manoil, C. and Beckwith, J. (1987) Proc. Natl. Acad. Sci. USA 84, 8525–8529.
[47] Miller, S.I., Kukral, A.M. and Mekalanos, J.J. (1989) Proc. Natl. Acad. Sci. USA 86, 5054–5058.
[48] Flügel, R.M., Rethwilm, A., Maurer, B. and Darai, G. (1987) EMBO J. 6, 2077–2084.
[49] Machamer, C.E. and Rose, J.K. (1987) J. Cell Biol. 105, 1205–1214.
[50] Holm, L., Saraste, M. and Wikström, M. (1987) EMBO J. 6, 2819–2823.
[51] Nakano, Y., Guerrero, I., Hidalgo, A., Taylor, A., Whittle, J.R.S.

and Ingham, P.W. (1989) Nature 341, 508–513.
[52] Pongs, O., Kecskemethy, N., Müller, R., Krah-Jentgens, I., Baumann, A., Kiltz, H.H., Canal, I., Llamazares, S. and Ferrus, A. (1988) EMBO J. 7, 1087–1096.
[53] Minegishi, T., Nakamura, K., Takakura, Y., Ibuki, Y. and Igarashi, M. (1991) Biochem. Biophys. Res. Commun. 175, 1125–1130.
[54] Mueckler, M., Caruso, C., Baldwin, S.A., Panico, M., Blench, I., Morris, H.R., Allard, W.J., Lienhard, G.E. and Lodish, H.F. (1985) Science 229, 941–945.
[55] Peralta, E.G., Winslow, J.W., Peterson, G.L., Smith, D.H., Ashkenazi, A., Ramachandran, J., Schimerlik, M.I. and Capon, D.J. (1987) Science 236, 600–605.
[56] Schofield, P.R., Darlison, M.G., Fujita, N., Burt, D.R., Stephenson, F.A., Rodriguez, H., Rhee, L.M., Ramachandran, J., Reale, V., Glencorse, T.A., Seeburg, P.H. and Barnard, E.A. (1987) Nature 328, 221–227.
[57] Connell, G.J. and Molday, R.S. (1990) Biochemistry 29, 4691–4698.
[58] Tanabe, T., Takeshima, H., Mikami, A., Flockerzi, V., Takahashi, H., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T. and Numa, S. (1987) Nature 328, 313–318.
[59] Hollmann, M., O'Shea-Greenfield, A., Rogers, S.W. and Heinemann, S. (1989) Nature 342, 643–648.
[60] Grenningloh, G., Rienitz, A., Schmitt, B., Methfessel, C., Zensen, M., Beyreuther, K., Gundelfinger, E.D. and Berz, H. (1987) Nature 328, 215–220.
[61] Namiki, M., Kitamura, M., Buczko, E. and Dufau, M.L. (1988) Biochem. Biophys. Res. Commun. 157, 705–712.
[62] Edwards, R.J., Murray, B.P., Singleton, A.M. and Boobis, A.R. (1991) Biochemistry 30, 71–76.