JOURNAL OF BACTERIOLOGY, May 1988, p. 2296–2300
0021-9193/88/052296-05$02.00/0

Vol. 170, No. 5

Statistical and Functional Analyses of Viral and Cellular Proteins with N-Terminal Amphipathic α-Helices with Large Hydrophobic Moments: Importance to Macromolecular Recognition and Organelle Targeting

MILTON H. SAIER, JR.,[1]* AND PETER McCALDON[2]

Department of Biology, University of California at San Diego, La Jolla, California 92093,[1] and European Molecular Biology Laboratory, D-6900 Heidelberg, Federal Republic of Germany[2]

Received 28 September 1987/Accepted 3 February 1988

A total of 1,911 proteins with N-terminal methionyl residues were computer screened for potential N-terminal α-helices with strong amphipathic character. By the criteria of D. Eisenberg (Annu. Rev. Biochem. 53:595–623, 1984), only 3.5% of nonplastid, nonviral proteins exhibited potential N-terminal α-helices, 18 residues in length, with hydrophobic moment values per amino acyl residue ($\langle\mu_H\rangle$) in excess of 0.4. By contrast, 10% of viral proteins exhibited corresponding $\langle\mu_H\rangle$ values in excess of 0.4. Of these viral proteins with known functions, 55% were found to interact functionally with nucleic acids, 30% were membrane-interacting proteins or their precursors, and 15% were structural proteins, primarily concerned with host cell interactions. These observations suggest that N-terminal amphipathic α-helices of viral proteins may (i) function in nucleic acid binding, (ii) facilitate membrane insertion, and (iii) promote host cell interactions. Analyses of potential amphipathic N-terminal α-helices of cellular proteins are also reported, and their significance to organellar or envelope targeting is discussed.

In a recent study, it was shown that eight of nine sequenced integral membrane permease proteins of the bacterial phosphotransferase system exhibit N-terminal amphipathic amino acyl sequences capable of forming α-helices with large hydrophobic moments (M. H. Saier, Jr., unpublished data; see references 8a and 12). Similar structures had been shown previously to be important to the import of eucaryotic, nuclear-encoded proteins into plastids such as mitochondria and chloroplasts (4–6, 10). It has also been postulated that a specific "nuclear targeting" sequence of different structure allows import of proteins into nuclei (3, 9). In spite of these facts and postulates, no detailed statistical analyses of N-terminal amphipathic α-helices have been published, and consequently the significance of the observations and postulates noted above remains questionable.

In this study, we screened nearly 2,000 methionine-terminal proteins for potential N-terminal, amphipathic α-helices. The results show that these structures occur with low frequency in several classes of proteins, but with significantly higher frequency in membrane and viral proteins. Most of the viral proteins of known function were shown to interact either with a nucleic acid or with membranes. A smaller number of these proteins mediate virus-host cell interactions. The presence of similar sequences in viral proteins of unknown or uncertain functions provides a clue as to their possible functions. The observations reported may be of significance with respect to a variety of subcellular assembly processes.

MATERIALS AND METHODS

About 3,500 proteins in the European Molecular Biology Laboratory (Heidelberg, Federal Republic of Germany) protein sequence data bank in January, 1987, were computer screened for amino-terminal sequences capable of forming amphipathic α-helices with hydrophobic moments of large magnitude. The systematic analysis of potential amphipathic helical N termini assumed that the 18 N-terminal residues were in a helical configuration regardless of the nature of the amino acyl residues within this segment of the protein. The presence of prolyl residues (common helix breakers) within and beyond the segment analyzed (to position 30) was subsequently noted and tabulated. The equation used to calculate the hydrophobic moment was

$$\mu_H = \left\{ \left[\sum_{n=1}^{N} H_n \sin\delta_n \right]^2 + \left[\sum_{n=1}^{N} H_n \cos\delta_n \right]^2 \right\}^{1/2}$$

where H is the hydrophobic moment of residue n, and δ is the angle in radians at which successive side chains emerge from the backbone. For the α-helical configuration, δ is 100°, and for the studies reported here, $N = 18$. Hydrophobicity values for individual amino acyl residues in proteins used for the computations were the consensus values previously reported (1). Initially, proteins of various classes or from various sources were statistically analyzed by the computer for N-terminal sequences with hydrophobic moments per residue ($\langle\mu_H\rangle$ values) in excess of 0.4, 0.5, 0.6, and 0.7 (Table 1). Subsequently, all proteins with N-terminal sequences exhibiting $\langle\mu_H\rangle$ values in excess of 0.4 were examined for function, and the relevant characteristics were summarized (see Tables 2, 3, and 4).

RESULTS

Nearly 4,000 proteins, about half of which possessed N-terminal methionyl residues, were computer screened for N-terminal amphipathic α-helices having large hydrophobic moments (Table 1). The N-terminal 18 amino acyl residues were examined by employing the consensus hydrophobicity values and computer programs described in Materials and Methods and reference 1. Methionine-terminal proteins showed the same statistical pattern as proteins with pro-

* Corresponding author.

TABLE 1. N-terminal, α-helical, hydrophobic moments of various classes of proteins[a]

Protein class	% of protein group having moments greater than:				Total no. of proteins
	0.4	0.5	0.6	0.7	
Methionine terminal proteins	5.7	1.7	0.4	0.1	1,911
Nonmethionine terminal proteins	5.8	2.1	0.4	0.2	1,631
Procaryotic proteins	4.6	1.2	0.2	0	417
Nonviral, nonmitochondrial, nonchloroplast proteins	3.5	0.3	0.1	0	1,052
Phage proteins	12.6	5.5	2.5	1.0	191
Animal viral proteins	8.2	2.6	0	0	466
Total viral proteins	10.0	3.2	0.7	0.3	710
Membrane proteins	10.2	4.0	0.3	0	75

[a] The proteins screened were those included within the Heidelberg (European Molecular Biology Laboratory) protein sequence data bank in January, 1987.

teolytically cleaved N termini (Table 1). Less than 6% of both groups of proteins exhibited N-terminal amphipathic helices with hydrophobic moments per residue ($\langle\mu_H\rangle$ values) greater than 0.4, and less than 2% possessed $\langle\mu_H\rangle$ values greater than 0.5. When viral, mitochondrial, and chloroplast proteins were deleted from the list of methionyl-terminal proteins, these values decreased to 3.5 and 0.3%, respectively (Table 1).

Every protein with an N-terminal methionine and a hydrophobic moment in excess of 0.4 was identified and examined. Only 13 cytoplasmic metabolic enzymes fell within this category (<1% of the sample screened), and of these, only 6 were from eucaryotic sources. The latter proteins included glyceraldehyde-3-phosphate dehydrogenase from *Thermus aquaticus* ($\langle\mu_H\rangle$ = 0.74; $\langle\bar{H}\rangle$ = 0.04; four positive charges [4+]; no negative charges in the leader sequence [0−]) and from human liver ($\langle\mu_H\rangle$ = 0.50; $\langle\bar{H}\rangle$ = 0.13; 4+; 0−) as well as yeast (*Saccharomyces cerevisiae*) thymidylate kinase ($\langle\mu_H\rangle$ = 0.42; $\langle\bar{H}\rangle$ = −0.02; 3+; 1−) and tryptophan synthase ($\langle\mu_H\rangle$ = 0.44; $\langle\bar{H}\rangle$ = −0.53; 4+; 2−). These proteins possessed leader sequences with an excess of positive charges over negative charges and lacking prolyl residues. *Escherichia coli* cytoplasmic proteins with similar N-terminal sequences included carbamoyl phosphate synthetase, small chain ($\langle\mu_H\rangle$ = 0.41; $\langle\bar{H}\rangle$ = +0.12; 3+; 2−; no prolyl residues), aspartokinase ($\langle\mu_H\rangle$ = 0.53; $\langle\bar{H}\rangle$ = +0.08; 3+; 1−; no prolyl residues), thymidylate synthase ($\langle\mu_H\rangle$ = 0.50; $\langle\bar{H}\rangle$ = −0.17; 3+; 3−; no prolyl residues), and acetohydroxyacid synthase ($\langle\mu_H\rangle$ = 0.41; $\langle\bar{H}\rangle$ = 0.18; 2+; 3−; residue 26 = proline). Several mitochondrial proteins exhibited similar properties, as expected (10), and chloroplast proteins with amphipathic leaders exhibiting either a net positive charge (e.g., pea cytochrome *f* precursor) or a net negative charge (e.g., pea ribulose bisphosphate carboxylase) were found. The N-terminal 18-residue sequence of the nuclear yeast mating protein a1 ($\langle\mu_H\rangle$ = 0.59; $\langle\bar{H}\rangle$ = 0.06) exhibited a net negative charge and no prolyl residues in its leader, but it lacked the nuclear targeting sequence proposed by Hall et al. (3).

Table 1 reveals that viral proteins and membrane proteins exhibited N-terminal amphipathic helices with large $\langle\mu_H\rangle$ values with increased frequency. Proteins from both bacterial and animal cell viruses exhibited this property. Since most of the nonplastid membrane-associated proteins were of viral origin, the values for these two classes are related.

Figure 1 reveals the distribution of proteins with N-terminal α-helical hydrophobic moments of various magnitudes. It can be seen that the group of nonviral, nonplastid proteins formed an approximately Gaussian distribution with a maximum at a $\langle\mu_H\rangle$ value of about 0.16 and a trailing edge at higher $\langle\mu_H\rangle$ values when the percent of the proteins falling within a particular range of $\langle\mu_H\rangle$ values was plotted versus these values. Viral proteins showed a similar distribution with the same maximum, but there was a pronounced shoulder at higher $\langle\mu_H\rangle$ values. When proteins from bacterial viruses were examined, this shoulder was even more pronounced (Fig. 1). While the group of membrane proteins examined also showed this tendency, the number of proteins analyzed was insufficient to give a statistically significant distribution (Table 1; data not shown).

Table 2 summarizes data for eucaryotic proteins or their precursors which are membrane associated during at least some stage of their biosynthesis. Except for the HLA histocompatibility antigen α-chain, blue-sensitive opsin, and the Na^+ channel protein, all such proteins were of viral origin. These proteins included the *gag* and *env* polyprotein precursors of several envelop viruses as well as the surface glycoproteins (or their precursors) of other related envelope viruses. Most of these leader sequences exhibit a net positive charge and a fairly hydrophilic character (Table 2). Excluding the integral membrane proteins of the bacterial phosphotransferase system (7, 8a, 11), the phycoerythrocyanine β-chain from the blue-green bacterium *Mastigocladus laminosus* (2) was the only integral membrane procaryotic protein screened with an N terminus exhibiting a large hydrophobic moment ($\langle\mu_H\rangle$ = 0.52).

By far the largest class of proteins with amphipathic α-helical N termini and known functions were nucleic acid-binding proteins. These proteins served a variety of func-

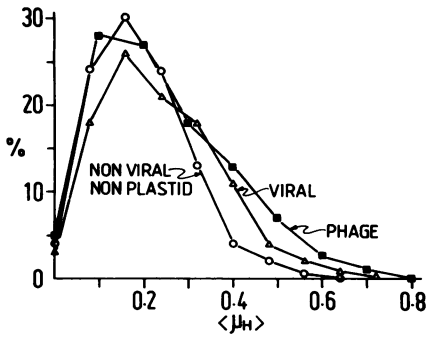

FIG. 1. Distribution of proteins with N-teminal, α-helical, hydrophobic moments, $\langle\mu_H\rangle$, of various magnitudes. The three classes of proteins shown are nonviral, nonplastid proteins (○); viral proteins (from animal, plant, and bacterial viruses) (△), and phage proteins (from bacterial viruses only) (■). The number of proteins in each group is given in Table 1.

TABLE 2. Eucaryotic membrane proteins with large N-terminal hydrophobic moments

Protein	Source[a]	Values for first 18 residues		No. of charges (first 18 residues)		Position(s) of prolyl residue(s) (first 30 residues)
		$\langle\mu_H\rangle$	$\langle\bar{H}\rangle$	+	−	
Animal						
HLA class II histocompatibility antigen, α-chain	Human	0.42	−0.85	7	2	15
Sodium channel protein	Electric eel	0.44	−0.35	5	1	18
Opsin, blue-sensitive	Human	0.45	−0.08	3	3	
Viral						
gag polyprotein (Precursor)	AIDS (simian retro-) virus SRV-1	0.41	−0.27	3	3	
	T-cell leukemia virus	0.60	−0.12	3	0	11, 13, 15, 16
	Avian spleen necrosis virus BH10	0.42	+0.23	2	1	12
env polyprotein (Precursor)	AIDS (simian retro-) virus HTLV-III	0.56	−0.29	6	1	
	Visna lentivirus	0.40	−0.45	4	3	19
	Bovine leukemia virus	0.48	−0.70	7	1	2, 13
Major surface glycoprotein G	Respiratory syncytial virus	0.44	−0.60	5	2	
Glycoprotein precursor	Pichinde arenavirus	0.49	−0.37	0	2	12
Latent membrane protein (LMP or the BNLF1 protein)	Epstein-Barr virus	0.53	−0.68	4	3	9, 10, 12, 15, 16, 18, 19

[a] AIDS, Acquired immunodeficiency syndrome; SRV-1, simian retrovirus type 1; HTLV-III, human T-cell lymphotropic virus type III.

TABLE 3. Proteins with N-terminal amphipathic α-helices with large hydrophobic moments (>0.4) which interact with nucleic acids

Protein (and function)	Source	Values for first 18 residues		No. of charges (first 18 residues)		Position(s) of prolyl residue(s) (first 30 residues)
		$\langle\mu_H\rangle$	$\langle\bar{H}\rangle$	+	−	
Phage						
RNA ligase	T4 (*E. coli*)	0.56	−0.29	3	3	
DNA ligase	T4 (*E. coli*)	0.41	+0.13	3	1	
Exodeoxyribonuclease	T7 (*E. coli*)	0.57	−0.08	2	3	9
Internal virion protein c	T7 (*E. coli*)	0.46	−0.05	3	1	13
Replication protein P	λ (*E. coli*)	0.44	−0.39	3	2	23
DNA-packaging protein	λ (*E. coli*)	0.47	−0.21	4	0	25, 27
Assembly (RNA packaging) protein	MS2 (*E. coli*)	0.50	−0.09	2	3	16
Regulatory protein ARC (transcriptional repressor)	P22 (*Salmonella typhimurium*)	0.55	−0.21	4	0	8, 15
Early protein gb6 (DNA replication)	Φ29 (*Bacillus subtilis*)	0.48	−0.51	4	1	
Virus (eucaryotic)						
DNA polymerase	Hepatitis virus (woodchuck)	0.59	−0.13	3	3	3, 25, 26
DNA polymerase	Hepatitis virus (squirrel)	0.50	−0.02	2	3	3, 25, 26
DNA polymerase	Herpes simplex virus type 1	0.40	+0.15	2	0	8, 11
DNA-binding protein (agnoprotein)	Simian virus 40	0.55	−0.43	6	0	
Phosphoprotein P	Measles virus	0.49	−0.22	5	4	23
Terminal protein	Adenovirus type 5	0.40	−0.13	4	2	6, 18
Terminal protein	Adenovirus type 7	0.46	−0.11	4	1	6, 18
Major capsid protein VP6	Bovine rotavirus, strain RF	0.56	−0.02	3	3	
Coat protein, precursor	Black beetle virus	0.43	+0.83	6	0	24, 26
Coat protein	Cauliflower mosaic virus	0.44	+0.03	2	2	
Coat protein	Alfalfa mosaic virus	0.41	−0.44	5	0	15
Cellular						
recF (recombination) protein	*E. coli*	0.52	−0.14	3	3	23
frn protein (transcriptional regulation)	*E. coli*	0.47	−0.06	5	1	3
Replication initiation protein	Plasmids R100 and R1	0.41	−0.22	2	1	15, 17, 21
Ribosomal protein S1	*E. coli*	0.49	+0.04	2	5	20
Ribosomal protein S33	Yeast (*S. cerevisiae*)	0.45	0	4	1	6
Ribosomal protein S1	Yeast (*S. cerevisiae*)	0.46	−0.33	7	1	22
Ribosomal protein S19	Tobacco chloroplast	0.48	−0.21	5	1	9
a1 Mating protein	Yeast (*S. cerevisiae*)	0.59	+0.06	1	3	
Heat shock protein 30	Frog (*Xenopus laevis*)	0.42	−0.65	5	2	
Heat shock protein 22	Fruit fly (*Drosophila melanogaster*)	0.62	−0.08	4	2	5, 18, 23

TABLE 4. Viral proteins of unknown or uncertain function with potential N-terminal α-helices with large hydrophobic moments ($\langle\mu_H\rangle \geq 0.4$)

Protein(s)	Virus
L2,[a] E2,[b,c] and E6[b]	Papillomavirus
BFRF1, BRRF1, BRLF1, and BGLF2 proteins	Epstein-Barr virus
BK protein	Herpes simplex virus
Early 28K,[b] 24K,[b] and 6.3K[b] proteins	Adenovirus type 7
A-106 protein	Adenovirus type 2
D287 and D551 proteins	Vaccinia virus
Nonstructural protein NS2[d]	Influenza A virus
S1 gene proteins[e]	Reovirus
VP6 (major inner capsid)[f] protein	Simian 11 rotavirus
B256 protein[g]	Cassava latent virus
Gene 278 protein[h]	Phage Pf3
Gene F protein[i]	Phage G4
Ea47 and Ea31 gene proteins[j]	Phage λ
nin region protein C-60	Phage λ
gene 1.1, 3.8,[j,k] 4.3,[l] 6.5, and 19.3[m] proteins	Phage T7

[a] The L2 protein shows homology with the β subunit of the proton-translocating ATPase of bacteria, chloroplasts, and mitochondria, but its function is unknown.

[b] These proteins are early proteins of DNA viruses, probably involved in transcriptional regulation. Consequently, they may be nucleic acid-binding proteins.

[c] The E2 protein shows homology with the c-*mos* protooncogene-encoded protein, but its function is unknown.

[d] The NS2 protein is likely to be a component of the influenza virus transcriptase and is therefore probably an RNA-binding protein.

[e] The S1 protein is an outer capsid hemagglutinin, probably required for host recognition and penetration. It may function in host-membrane association.

[f] The VP6 protein is probably a nucleic acid-binding protein.

[g] The B256 protein is required for systemic spread of the virus throughout the host plant and may therefore be involved in host recognition as well as membrane association.

[h] The gene 278 protein is possibly involved in the DNA replication process.

[i] The gene F protein is a virion structural protein, probably a DNA-binding protein.

[j] These two proteins may be DNA endonucleases.

[k] The gene 3.8 protein appears to be a host-recognition protein involved in cell lysis.

[l] The gene 4.3 protein may insert into the membrane of the host cell and promote K^+ efflux.

[m] The gene 19.3 protein is encoded by a gene in the region of the T7 genome involved in DNA packaging. It may therefore be a DNA-binding protein.

tions (Table 3). While the vast majority of these proteins were of viral origin, cellular proteins involved in recombination, transcriptional regulation, DNA replication, and RNA translation were also found.

Table 4 summarizes data for viral proteins (some of which are still hypothetical) of unknown or poorly defined functions which were found to possess large N-terminal, α-helical hydrophobic moments. Several of these leader sequences showed highly distinctive properties. For example, the N terminus of the gene 1.1 protein of phage T7 exhibited the largest $\langle\mu_H\rangle$ value of any protein tested (0.77). The N-terminal 18 residues are shown in helical wheel depiction in Fig. 2A. The polarity of hydrophobicity, as well as the unequal charge distribution, is apparent. For the gene 1.1 protein of phage T7 and many of the other viral proteins listed in Table 4, a function involving an interaction with nucleic acids seems likely.

A smaller class of phage proteins with large N-terminal hydrophobic moments was found to serve a structural role, usually functioning in host interactions. These proteins included the minor tail proteins H, T, and Z of phage λ ($\langle\mu_H\rangle$ values of 0.43, 0.54, and 0.62, respectively) as well as the host specificity protein B of phage T7. The last-mentioned protein showed a most distinctive N-terminal amphipathic α-helix with 10 uninterrupted lysyl residues on one side and 11 neutral residues, the majority of which are valyl residues, on the other side. The only two semipolar residues in the helix (Ser-15 and Thr-13) compose the boundary between the hydrophobic and hydrophilic sides of the helix (Fig. 2B). The hydrophobic moment for this helix ($\langle\mu_H\rangle = 0.76$) and the average hydrophobicity value ($\langle\bar{H}\rangle = -0.42$) clearly reflect a highly directed evolutionary process.

DISCUSSION

The data presented in this paper provide a statistical analysis and functional evaluation of the N-terminal, amphipathic, potentially α-helical sequences in a large number of proteins. We have found that when mitochondrial, chloroplast, and bacterial phosphotransferase proteins are excluded from the list of proteins screened, very few natural cellular protein constituents of either procaryotes or eucaryotes possess extended N-terminal amphipathic α-helical structures with large hydrophobic moments. When the N-terminal 18 residues are analyzed as described in Materials and Methods (1), only about 1 in 30 such proteins have $\langle\mu_H\rangle$ values in excess of 0.4. A few nonplastid, soluble and membrane constituents were found to possess such amphipathic structures, however, and most of these sequences bore an excess of positive charge while lacking α-helix-disrupting residues such as prolyl and pairs of adjacent glycyl residues. A prolyl residue or two adjacent glycyl residues terminate the amphipathic α-helical N termini of each of the permeases of the bacterial phosphotransferase system (Saier, unpublished data; see references 8a and 12).

Amphipathic leader sequences are believed to be both necessary and, when appropriately localized in the protein, sufficient for import of the protein into mitochondria (3–5, 8a, 9). However, since some of the proteins found to possess these N-terminal sequences were not mitochondrial constituents, it is clear that the presence of such a structure alone is insufficient to provide all of the necessary targeting information. Not all of the information presented in this report is consistent with current postulates regarding chloroplast and nuclear targeting (3, 8a). Further investigations will be required to establish the nature and extent of targeting information carried by a eucaryotic protein.

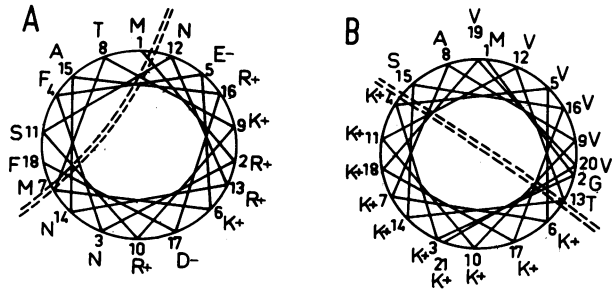

FIG. 2. Helical wheel projections of (A) the N-terminal 18 amino acyl residues of the gene 1.1 protein of phage T7 and (B) the N-terminal 21 amino acyl residues of the host specificity protein B of phage T7. Numbers indicate the numbers of the amino acyl residues from the N terminus of the protein. The capital letter adjacent to the number is the one-letter abbreviation for the amino acyl residue at that position (8). The charge of a basic residue, arginine (R) or lysine (K), is indicated by a plus (+), while the charge of an acidic residue, aspartate (D) or glutamate (E), is indicated by a minus (−).

A protein which does not possess an N-terminal 18-aminoacyl segment with a hydrophobic moment in excess of 0.4 may nevertheless possess an N-terminal amphipathic structure, which allows it to be targeted to the mitochondria of eucaryotes or the envelope fraction of procaryotes, for several reasons. (i) An amphipathic sequence of only 12 residues may be sufficient to target eucaryotic proteins to mitochondria (4) and bacterial proteins to the envelope (Y. Yamada, M. Yamada, and M. H. Saier, Jr., unpublished data). (ii) Although all recognized amphipathic leader sequences of mitochondrial proteins are near the N termini, they may be displaced several residues from the N terminus (9, 10). Interestingly, this was never observed for the eight sequenced permeases of the bacterial phosphotransferase system (7, 8a, 12; Saier, unpublished data). (iii) Amphipathic leaders may not need to exhibit large $\langle \mu_H \rangle$ values to target a protein to the membrane. Thus, the amphipathic, N-terminal leader sequences of bacteriorhodopsin and halorhodopsin (11) exhibit magnitudes lower than those we have screened. It should be noted, however, that in these last-mentioned cases, the function of these leader sequences has not yet been investigated.

A most striking observation which resulted from the reported analyses concerned the frequency, relative to cellular proteins, with which viral proteins exhibited strongly amphipathic, α-helical N termini (Fig. 1 and Table 1). That these structures are of functional significance is suggested by the analysis of proteins found to possess these sequences. Thus, DNA- and RNA-binding proteins constituted the largest class (55%; Table 3), membrane proteins or precursor proteins which associate with membranes made up a very significant but smaller group (30%; Table 2), and structural constituents of bacteriophage which function in host interactions made up a still smaller class of proteins (15%). It will be interesting to determine what the frequency of internal and C-terminal amphipathic α-helices is relative to that of the N-terminal amphipathic α-helices analyzed here when the different functional classes of proteins are considered. These studies are currently in progress.

Many viral proteins of unknown or uncertain function were found to possess N-terminal amphipathic sequences (Table 4). The analyses reported suggest that the principal roles of N-terminal amphipathic α-helices in proteins are (i) facilitation of interactions with nucleic acids, (ii) association of proteins with biological membranes, and (iii) recognition of host cell constituents. All of these functions fall within the category of macromolecular recognition. These observations provide a guide for studies aimed at the elucidation of the functions of proteins with N-terminal amphipathic structures such as those listed in Table 4.

ACKNOWLEDGMENTS

We thank the Alexander von Humboldt Foundation of W. Germany for a senior scientist award to M.H.S. which rendered these studies possible. M.H.S. also wishes to acknowledge the generous hospitality of Gerhart Drews, University of Freiburg. This work was supported by Public Health Service grants 5 R01-AI-21702 and 2-R01-AI 14176 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Eisenberg, D.** 1984. Three-dimensional structure of membrane and surface proteins. Annu. Rev. Biochem. **53**:595–623.
2. **Fuglistaller, P., F. Suter, and H. Zuber.** 1983. The complete amino-acid sequence of both subunits of phycoerythrocyanine from the thermophilic cyanobacterium *Mastigocladus laminosus*. Hoppe-Seyler's Z. Physiol. Chem. **364**:691–712.
3. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of *E. coli* β-galactosidase to the nuclei of yeast. Cell **36**:1057–1065.
4. **Hurt, E. C., B. Pesold-Hurt, K. Suda, W. Opplinger, and G. Schatz.** 1985. The first twelve amino acids (less than half of the pre-sequences) of an imported mitochondrial protein can direct mouse cytosolic dihydrofolate reductase into the yeast mitochondrial matrix. EMBO J. **4**:2061–2068.
5. **Hurt, E. C., N. Soltanifar, M. Goldschmidt-Clermont, J.-D. Rochaix, and G. Schatz.** 1986. The cleavable pre-sequence of an imported chloroplast protein directs attached polypeptides into yeast mitochondria. EMBO J. **5**:1343–1350.
6. **Roise, D., S. J. Horvath, J. M. Tomich, J. H. Richards, and G. Schatz.** 1980. A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphipathic helix and perturb natural and artificial phospholipid bilayers. EMBO J. **5**:1327–1334.
7. **Saier, M. H., Jr.** 1985. Mechanisms and regulation of carbohydrate transport in bacteria. Academic Press, Inc., Orlando, Fla.
8. **Saier, M. H., Jr.** 1987. Enzymes in metabolic pathways: a comparative study of mechanism structure, evolution and control, p. 2–6. Harper & Row, Publishers, Inc., New York.

8a. **Saier, M. H., Jr., M. Yamada, B. Erni, K. Suda, J. Lengeler, R. Ebner, P. Argos, B. Rak, K. Schnetz, C. A. Lee, G. C. Stewart, F. Breidt, Jr., E. G. Waygood, K. G. Peri, and R. F. Doolittle.** 1988. Sugar permeases of the bacterial phosphoenolpyruvate-dependent phosphotransferase system: sequence comparisons. FASEB J. **2**:199–208.

9. **Schekman, R.** 1985. Protein localization and membrane traffic in yeast. Annu.Rev. Cell Biol. **1**:115–143.
10. **von Heijne, G.** 1986. Mitochondrial targeting sequences may form amphiphilic helices. EMBO J. **5**:1335–1342.
11. **von Heijne, G., and J. P. Segrest.** 1986. The leader peptides from bacteriorhodopsin and halorhodopsin are potential membrane-spanning amphipathic helices. FEBS Lett. **213**:238–240.
12. **Yamada, M., and M. H. Saier, Jr.** 1987. Glucitol-specific enzymes of the phosphotransferase system in *Escherichia coli*. Nucleotide sequence of the *gut* operon. J. Biol. Chem. **262**: 5455–5463.