5. Conclusions

The present analysis has shown that the nature of variation of amplitude and phase of the semidiurnal component of neutral wind are predominantly different in the successive height intervals of 80–90 km and 90–100 km.

1. The amplitudes of both EW and NS components have predominantly negative gradients from 80–90 km, the same have positive gradients from 90–100 km.

2. The height gradient in amplitude of the EW component from 90–100 km is found to exhibit a semi-annual variation and the same gradient is found to exhibit a similar nature, to some extent, but with peaks occurring one month in advance between 80–90 km height interval.

3. The height gradient in amplitude of the NS semidiurnal component from 90-100 km is found to exhibit oscillations with a period of 3 to 4 months, and the same component is found to exhibit similar oscillation between 80 and 90 km, but with much smaller amplitude in almost all months of the year.

4. The height gradient and phase of the NS semidiurnal component between 80 and 90 km is found to fluctuate between positive and negative values as the same in 90–100 km interval, but these fluctuations are strikingly out of phase with the fluctuations in the 90–100 km interval.

5. Almost similar behaviour is exhibited by the gradients EW of phase between 80 and 90 km and 90 and 100 km. These fluctuations in phase are expected to be due to the coupling between main semidiurnal mode and higher order modes.

Acknowledgements

The authors are thankful to Dr. H. G. Muller, University of Sheffield, England, for passing on the Meteor Wind Data of Adelaide station. One of the authors (PCSD) wishes to thank the University Grants Commission, New Delhi, for financial support.

1. Greenhow, J. S., *J. Atoms. Terr. Phys.*, 1952, **2**, 282.
2. —, *Phil. Mag.*, 1954, **45**, 471.
3. — and Neufeld, E. L., *Ibid.*, 1955, **46**, 549.
4. — and —, *Proc. Phys. Soc.*, 1959, **74**, 1.
5. Müller, H. G., *Planet. Space Sci.*, 1966, **14**, 1253.
6. Greenhow, J. S. and Neufeld, E. L., *Phil. Mag.*, 1956, **1**, 1157.
7. Müller, H. G., *J. Atmos. Terr. Phys.*, 1968, **30**, 701.
8. Glass, M. and Spizzichino, A., *Ibid.*, 1974, **36**, 1825.
9. Fraser, G. J., *Ibid.*, 1968, **30**, 707.
10. Elford, W. G., *IAMAP/IAPSO Combined First Special Assembly, Melbourne*, 1974, p. 624.
11. Belmont, A. D. and Dartt, D. G., *J. Geophys. Res.*, 1973, **78**, 6373.
12. Wilkinson, P. J. and Baggaley, W. J., *Planet. Space Sci.*, 1975, **23**, 503.

A COMMON FACTOR IN *IN VIVO* SYNTHESISED POLYPEPTIDES

A. S. KOLASKAR

School of Life Sciences, University of Hyderabad, Hyderabad 500 001

AND

K. V. SOMAN

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012

FROM early days many workers have looked at the composition of the protein and tried to relate it with the three-dimensional structure[1]. However, it has been well established that the three-dimensional structure is the resultant of amino acid sequence[2]. Furthermore, initial studies have indicated that the oil-drop model proposed by Kauzmann[3] based on hydrophobic interactions can explain folding of polypeptide chain. However, the recent studies of Chothia[4] have indicated that the principle of close packing of different amino acids governs the three-dimensional structure of proteins. Therefore, in order to get some insight into the role of hydrophobic interactions, the importance of which cannot be underrated, the composition of large number of proteins have been analysed.

Our studies on number of proteins, enzymes and polypeptides have shown that the total number of non-hydrophobic residues are directly related to the total number of residues in a given *in vivo* synthesised polypeptide chain indicating that the stability of a protein may be governed by its composition, mainly the number of hydrophobic residues.

The amino acids without any polar side groups, namely, Ala, Ile, Leu, Met, Phe, Pro, Trp and Val are considered to be hydrophobic. Remaining twelve amino acids have been termed as non-hydrophobic[5],

The forty different proteins which have been considered for analysis vary greatly in their properties as well as in the number of amino acid residues. As can be seen from Table I, the total number of amino acid residues vary from 48 to 944. In this set those proteins for which the crystal structure is known have been included mainly to see the relation, if it exists, between number of hydrophobic residues and the three-dimensional structure of proteins. In Table I, we have listed the number of non-hydrophobic residues

TABLE I

Total number of amino acid residues and the number of nonhydrophobic residues in proteins

Sl. No.	Protein	Source	No. of residues	
			Total	nonhydrophobic
1.	Posterior pit. peptide	Bovine	48	32
2.	Rubredoxin	*Clostridium pasteurianum*	54	40
3.	Trypsin Inhibitor	BP	58	38
4.	Basic plasma protein	Human	81	52
5.	Proinsulin	Porcine pancreas	84	32
6.	Cytochrome b5	Calf liver	85	58
7.	HIPIP	Chromatium	85	45
8.	β-Lipoprotein	Sheep	90	57
9.	Ferredoxin	Alfalafa	97	63
10.	Histone IV	Calf thymus	102	68
11.	RNase T1	*Aspergillus oxyzae*	104	75
12.	Parvalbumin (Myogen)	Carp muscle	108	59
13.	Coat protein	Bacteriophage fd.	111	64
14.	Cytochrome c2	*Rhodospirillum rubrum*	112	68
15.	RNase S	Bovine pancreas	124	87
16.	Lysozyme	HEW	129	84
17.	Flavodoxin	Clostridium	138	84
18.	Hemoglobin α-chain	Horse	141	77
19.	Nuclease	Staphylococcal	142	90
20.	Hemoglobin β-chain	Horse	146	79
21.	Hemoglobin	Sea lamprey	148	77
22.	Myoglobin	Sperm whale	153	87
23.	Coat protein	TMV-strain vulgare	158	90
24.	Kunitz inhibitor	Glycine mase soyabean	181	107
25.	β1-Glycoprotein	Human plasma	187	122
26.	Growth hormone	Human	188	116
27.	Bence Jones protein	ROY	214	139
28.	Trypsinogen	Bovine pancreas	229	151
29.	Conconavaline A	Jaekbean	237	141
30.	α-Chymotrypsin	Bovine	241	142
31.	Chymotrypsinogen A	Bovine	245	146
32.	Carbonic anhydrase B	Horse erythrocytes	265	164
33.	Deoxyribonuclease	Bovine pancreas	268	155
34.	Subtilisin BPN'	*Bacillus amylolique faciens*	275	156
35.	Carboxy peptidase A	Bovine	307	190
36.	Thermolysin	*Bacillus thermoproteolyticus*	316	208
37.	M4 Apo LDH	Dogfish	329	188
38.	Glyceraldehydre-3-ph.	Lobster	333	187
39.	Serum albumin	Bovine	566	351
40.	DNA Polymerase	T4-infected *E. coil*	944	590

which are usually more in number as compared to hydrophobic residues and the total number of residues in each protein. If N_{tot} and N_{nhb} are respectively the numbers of the total and non-hydrophobic residues, in a given protein, then as given in Fig. 1, there exists a relation

$$N_{nhb} = 0{\cdot}619\, N_{tot} - 2{\cdot}164 \qquad (1)$$

TABLE II

Application of the derived relation to other proteins, enzymes and polypeptides synthesised in vivo

Sl. No.	Protein	Source	N_{tot}	Nonhydrophobic		Error (%)
				Actual	Calc.	
1.	Triose phosphate isomerase	Rabbit muscle	248	142	151	6·34
2.	Elastase	Porcine	240	149	146	2·01
3.	α-Lytic Protease		198	123	120	2·44
4.	Papain		212	138	129	6·52
5.	Subtilisin	Carlsberg	274	156	167	7·05
6.	α-Lactalbumin	Human	274	79	74	6·33
7.	Endolysin	λ-phage	157	98	95	3·06
8.	Asp. aminotransferase	Pig heart	412	232	253	9·05
9.	Phospholipase A	Bee vemom	129	92	78	16·13
10.	Acyl carrier protein	*E. coli*	77	47	45	4·26
11.	Asp. transcarbamylase	*E. coli*	152	89	91	2·25
12.	Penicillinase	*S. aureus*	257	163	157	3·68
13.	Aldolase	Rabbit muscle	361	214	221	3·27
14.	Pepsin	Porcine	327	199	200	0·50
15.	Adenylate kinase	Porcine	194	125	118	5·60
16.	Deoxyribonuclease A	Bovine	257	150	157	4·67
17.	Dihydrofolate Reductase	*E. coli*	156	84	94	13·00
18.	Asparaginase	*E. coli*	321	194	197	1·55
19.	Protease A	*Strepto. griseus*	182	121	110	9·09
20.	Dismutase	Bovine	151	99	91	8·08
21.	Ferredoxin	Cl. pas.	55	32	32	0·0
22.	α-Neurotoxin	Egyptian cobravenom	61	51	36	29·4
23.	Azurin	Pseudomonas	128	80	77	3·8
24.	Retinoi binding Retino protein	Human	176	109	107	1·8
25.	Kallikrein inhibitor 2	Var. Dan.	204	124	124	0·0
26.	Kallikrein inhibitor 1	Var. Dan.	209	136	127	6·6
27.	Pepsin inhibitor	Asc. lumbri	290	161	177	9·9
28.	Ovo inhibitor	*Gallus gallus*	402	287	247	13·9
29.	ACTH		39	23	21	4·0
30.	Insulin		51	33	30	10·0
31.	Glutamate dehydrogenase	Bovine	500	306	308	0·7
32.	Actin	Rabbit muscle	515	307	317	3·3
33.	Myosin	Bovine	885	559	546	2·3
34.	α-trypsin inhibitor	Human	1224	739	756	2·3
35.	α2-Macroglobulin	Human	6463	3846	4001	4·0
36.	Cro regulatory protein	Phage	66	40	39	2·5
37.	S-Sulphofibrinogen	Human	2906	1970	1787	8·78
38.	Flagellin	SW 1061	382	239	234	2·08
39.	Keratin-SCMKB-IIIB2	Wool	97	61	58	4·91
40.	Collagen	Chicken	1052	632	624	2·0

Data for Table I and Table II has been collected from refs. 8 and 9.

The correlation coefficient value 0·996 which has been obtained from least square line fit indicates that the fit is very good.

The linear relation, Eqn. 1, was applied to many other enzymes, proteins and polypeptides. As can be seen from Table II, the actual number of non-hydrophobic residues in a given protein when compared with those obtained using Eqn. 1, are in error by less than 10% in most cases. It should be noted that the same relation holds good even for fibrous proteins such as Collagen, Keratin, Flagellin and S-Sulphofibronogen. This relation also holds good for inactive hormones like pro-insulin or active hormones like insulin and ACTH, thus, indicating that this relation has little relevance to the three-dimensional structure or activity of the polypeptide chain. However, analysis of crystal structure data of twenty-two globular proteins have showed that the number of reversals of a chain or turns in a protein can be expressed in terms of number of non-hydrophobic residues in that protein by the linear relation

$$T = 0{\cdot}2044\ N_{nhb} + 6{\cdot}03 \qquad (2)$$

where T is number of turns in a protein.

A similar relation has been recently obtained by Rose and Wetlaufer[6] for total number of residues in a protein and number of turns. Thus, it is very difficult to state the biological significance of this relation, but it may throw some light on factors which govern the thermodynamical stability of polypeptide chain in *in vivo* conditions. The indication of this fact comes from the analysis on the molecule ACTH. It has been shown that only the first 24 residues of this molecule give 100% activity[7]. However, *in vivo* secreted ACTH always contains 39 residues. When the analysis was carried out for the first 24 residues of this molecule, it was found that the predicted number disagrees with the actual number of Hydrophobic residues by about 20%, while, the results were in very good agreement for the complete ACTH molecule (See Table II). Secondly, for both the molecules pro-insulin, and insulin which are synthesised using cell machinery, the Eqn. 1 gives the results which are in very good agreement with the observed composition.

It can be readily seen that the ratio of twelve non-hydrophobic residues to total twenty amino acid residues which commonly occur in *in vivo* synthesised polypeptide chains is 0·6, nearly equal to the slope of the line showed in Fig. 1. Similarly, if one calculates the ratio of number of codons for hydrophobic

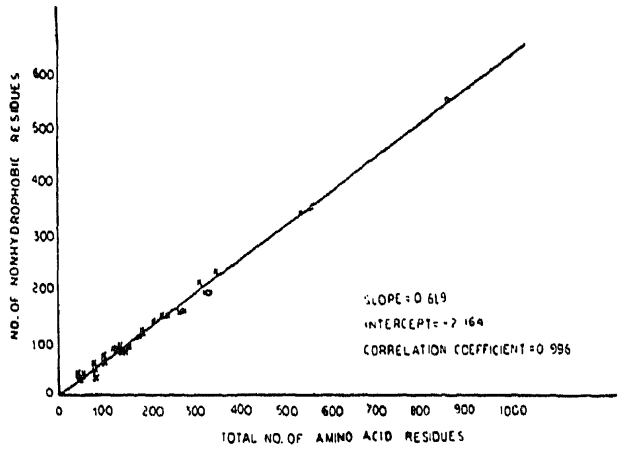

FIG. 1

residues mentioned above to total number of codons one gets the number 0·3906. This means, percentage of codons for nonhydrophobic residues is 61% exactly the same as the slope for the least square line which has been obtained from the analysis of protein data. This, indicates that the relation obtained above has definite biological significance though not understood at this level. Our studies in this direction are in progress.

One fact, which seems to be quite important and comes out from this analysis is that there is a common feature among almost all *in vivo* synthesised polypeptide chains, namely, the definite relation between the number of hydrophobic residues and total number of residues.

1. Fisher, H. F., *Proc. Natnl. Acad. Sci. (USA)*, 1964, **51**, 1285.
2. Anfinsen, C. B., *Science*, 1973, **181**, 223.
3. Kauzmann, W., *Adv. Protein Chem.*, 1959, **14**, 1.
4. Chothia, C. H., *Nature*, 1975, **254**, 304.
5. Lenhinger, A. L., *Biochemistry*, Worth Publ., 1970.
6. Rose, D. G. and Wetlaufer, D. B., *Nature*, 1977, **268**, 769.
7. McKerns, W. (Ed.) *Functions of the Adrenal Cortex*, Appleton-Century-Crofts, New York (1968).
8. Fasman, G. D. (Ed.) *Handbook of Biochemistry and Molecular Biology*, Proteins III, CRC Press, 1976.
9. Croft, A. R., *Handbook of Protein Sequences*, Oxford Press, 1976.