Vol. 156, No. 1, 1988
October 14, 1988

CALCIMEDIN, CALELECTRIN: CORRELATION OF RELATEDNESS

Pamela B. Moore

The Rockefeller University
New York, New York 10021-6399

Received August 26, 1988

A statistical method has been used to compare the amino acid compositions of several calelectrin proteins and the 67k calcimedin protein. The validity of the method for determining protein relatedness and ancestry has been established with many different proteins. Torpedo calelectrin, p68 brain calelectrin, 67,000 dalton brain calelectrin, 67,000 bovine aorta protein, 32k lipocortin (and probably 67R lung protein) share considerable homology with each other and with calpactin I or lipocortin II, the pp60^{src} tyrosine kinase substrate. The 67k calcimedin appears to be unrelated to these several proteins, in agreement with biochemical evidence, although limited homology may still be present. The lung 67E protein, calregulin and calmodulin, are unrelated to any of the other proteins.

Numerous calcium-binding proteins which bind to synthetic hydrophobic ligands or to biomembranes in the presence of micromolar calcium have been described during the past several years. These proteins include the calcimedins (1), brain calelectrin (2), the chromobindins (3), the 68,000 dalton lymphocyte protein (4) and calregulin or CAB63 (5). Little is known about the relationships or functional homologies of these calcium-binding proteins since well documented comparative data are not available. Statements in the literature have suggested that all these proteins, derived from different sources, are identical with each other (6,7,8) based solely on crossreactivity of antibody. Biochemical data obtained with several of these proteins does not support this identity.

Although sequences of two of these proteins have now been determined (9,10), the amino acid compositions are known for eleven of these proteins. The well established SΔQ method of Marchalonis and Weltman (11,12) allows the determination of protein relatedness based on amino acid composition (11-16). The method calculates SΔQ values for any two proteins being compared. SΔQ is a measure of the diversity of the two proteins. An SΔQ value <50 suggests primary structure homology (11); the stronger the relatedness, the lower the SΔQ value. When analyses of the same protein performed by different laboratories were compared, the maximum SΔQ obtained was 4. Unrelated proteins had SΔQ values >50; and values >50 implied little or no relatedness (11). The validity of this method has been demonstrated with a wide range of proteins, including immunoglobulin μ chains (13), lysozymes (14), membrane glycoproteins and transplantation antigens (15), carotenoid containing lipoglycoproteins (16), and 177 miscellaneous proteins (11,12).

0006-291X/88 $1.50

Methods

The relatedness coefficient, SΔQ, is a defined as the sum of the squares of each amino acid difference, or $\Sigma (X_{i,j} - X_{k,j})^2$, where X_j is the content of the amino acid j in each of the two proteins (i and k) being compared.

Results and Discussion

Eleven calcium-binding proteins are compared by the SΔQ criterion. Eight of these proteins are about the same size (65,000 daltons to 70,000 daltons); several papers have suggested that these proteins are identical (6,7,8). Three additional proteins, Torpedo calelectrin (M_r = 34,000), 32k lipocortin (M_r = 32,000) and calmodulin (M_r = 17,000), were also compared. The relatedness coefficients were calculated from the amino acid composition literature values reported for these eleven proteins. Table I is a compilation of the calculated amino acid compositions in mole percent used for the subsequent relatedness coefficient (SΔQ) calculations. Table II is a compilation of the calculated coefficients (SΔQ) between these eleven proteins taken two at a time.

TABLE I

AMINO ACID COMPOSITION COMPARISON (Mol %)

Residue	67k Calcimedin	p68 Calelectrin	67,000 Calelectrin	Torpedo Calelectrin	32k Lipocortin	67k Bovine Aorta	Lung 67R	Lung 67E	68,000 Lymphocyte	CAB 63	Brain Calmodulin
Asx	9.5	10.6	10.5	12.8	9.9	11.2	9.6	1.1	12.2	18.6	15.6
Thr	4.6	5.8	5.3	6.2	6.0	5,7	5.1	6.0	6.3	3.7	7.0
Ser	8.7	6.3	8.3	6.9	8.3	5.8	8.9	12.7	6.2	3.3	2.6
Glx	15.7	12.9	13.2	13.5	13.8	13.0	11.8	23.3	14.8	22.0	20.9
Gly	11.9	6.8	9.0	7.3	9.3	8.5	15.0	13.4	5.4	4.3	4.4
Ala	9.8	8.7	8.3	8.8	8.2	9.0	7.0	8.0	6.3	3.5	5.2
Val	3.8	5.0	2.6	5.5	4.0	4.0	3.8	4.6	2.3	4.2	4.6
Met	1.8	2.5	2.6	1.5	1.2	2.3	2.5	0.5	2.9	1.2	7.3
Ile	2.8	6.4	4.9	5.5	4.5	6.2	5.3	2.9	2.7	4.9	5.2
Leu	6.1	11.9	9.8	9.1	10.7	9.5	9.1	5.2	8.3	4.9	6.4
Tyr	2.6	0.49	2.6	2.9	3.8	3.4	4.1	2.0	5.6	5.4	1.8
Phe	3.7	3.6	3.0	2.9	4.0	3.1	3.3	2.3	5.1	6.3	6.5
His	4.3	1.8	1.5	1.8	1.5	1.4	2.1	2.6	5.7	2.5	0.8
Lys	9.0	8.0	9.0	6.6	7.6	7.9	6.8	3.9	9.8	11.8	5.7
Arg	2.3	6.2	5.6	4.4	5.7	5.2	5.6	1.8	6.3	3.4	6.0

Amino acid compositions in mole %: 67k calcimedin [1], p68 calelectrin [21], 67,000 calelectrin and Torpedo calelectrin [2], lipocortin [26], bovine aorta protein [20], lung 67R and 67E proteins [23], 68,000 lymphocyte protein [4], CAB 63 [5], and calmodulin [24]. Proline, tryptophan and cysteine contents were not consistently reported for the eleven proteins and were deleted from the calculations.

TABLE II

COMPARISON OF CORRELATION COEFFICIENTS (SΔQ)

	67k Calcimedin	p68 Calelectrin	67,000 Calelectrin	Torpedo Calelectrin	32k Lipocortin	67k Bovine Aorta	Lung 67R	Lung 67E	68,000 Lymphocyte	CAB 63	Brain Calmodulin
67k Calcimedin	--	119	58	80	63	74	72	186	110	283	267
p68 Calelectrin	119	--	[28]	[27]	[31]	[20]	105	386	95	301	193
67,000 Calelectrin	58	[28]	--	[29]	[11]	[13]	50	286	67	277	214
Torpedo Calelectrin	80	[27]	[29]	--	[26]	[12]	88	344	74	228	163
32k Lipocortin	63	[31]	[11]	[26]	--	[17]	[46]	274	72	286	221
67k Bovine Aorta	74	[20]	[13]	[12]	[17]	--	64	345	71	252	181
Lung 67R	72	105	50	88	[46]	64	--	277	156	405	327
Lung 67E	186	386	286	344	274	345	277	--	411	602	503
68,000 Lymphocyte	110	95	67	74	72	71	156	411	--	164	156
CAB 63	283	301	277	228	286	252	405	602	164	--	124
Brain Calmodulin	267	193	214	163	221	181	327	503	156	124	--

SΔQ analysis of the twelve calcium-binding proteins. SΔQ is a measure of protein relatedness [9,10]. Values less than 50 have been outlined.

The apparently related proteins (SΔQ <50) include Torpedo calelectrin, lipocortin, the 67,000 and p68 brain calelectrins, and the bovine aorta 67,000 dalton protein. The lung 67R protein appears separate from this group. However, differences in only one amino acid (glycine) accounts for 60-70% of the SΔQ values, suggesting the protein does belong to this group of proteins. Several of these proteins have also been reported to share a 17 amino acid consensus sequence with p36 calpactin I (17), the $pp60^{src}$ tyrosine kinase substrate (18,19). The SΔQ values between these proteins and calpactin I ranged between 16 and 33.

Proteins with SΔQ values between 50 and 100 include the 67k calcimedin (1) and the 68,000 dalton lymphocyte protein (4). These proteins do not appear to belong to the lipocortin/ calelectrin family above. SΔQ values suggest that these two proteins are unrelated to each other and would share only limited homology, if any, with the first group. Despite the apparent non-identity of these proteins based on analysis of the reported amino acid compositions, the lymphocyte 68,000 Da protein has now been reported to be identical to the 67,000 Da calelectrin protein (9,10). Analysis of the compositions predicted from the cDNA sequences of calelectrin and the lymphocyte protein gave an SΔQ of 0.5% (Table III; the 67k calcimedin is included for comparison). The

TABLE III

COMPARISON OF CALELECTRIN SEQUENCE COMPOSITIONS
WITH 67k CALCIMEDIN AMINO ACID COMPOSITION

Residue	Human Calelectrin	Lymphocyte Calelectrin	Smooth Muscle Calcimedin
Asx	10.9	10.8	9.5
Thr	5.1	5.2	4.6
Ser	4.7	4.8	8.7
Glx	13.7	13.4	15.7
Pro	2.0	2.0	2.9
Gly	3.6	3.8	11.9
Ala	6.0	6.1	9.8
Val	3.5	3.4	3.8
Met	3.4	3.6	1.8
Cys ½	0.8	0.8	nd
Ile	7.0	7.2	2.8
Leu	10.0	10.1	6.1
Tyr	4.8	4.8	2.6
Phe	4.7	4.4	3.7
Trp	0.5	0.5	0.8
His	2.3	2.1	4.3
Lys	8.7	8.8	9.0
Arg	8.2	7.9	2.3

nd=not determined

Comparison of the amino acid sequence compositions for the human calelectrin and the lymphocyte proteins. The amino acid compositions of the 67k calcimedin is included for comparison.

lymphocyte protein appears to have a true molecular size ranging between 80,000 and 85,000 daltons, rather that the 67,000 Da, estimated from SDS-polyacrylamide gel electrophoresis (PAGE). The sequence data suggest this increased size is accounted for by the linker regions between the repeat residues (9).

The apparent discrepancy between the published amino acid analyses for the 67k Da brain and the p68 lymphocyte calelectrins and the published sequence analyses for the human and lymphocyte calelectrins, may be partially resolved by the following: small amounts of the 67k calcimedin derived from blood vessels may be present in the calelectrin preparations and go undetected on 10-12% SDS-polyacrylamide gels which do not resolve the two proteins. Antibody to the 67k calcimedin does not recognize brain calelectrin (20) and does not crossreact with the lymphocyte protein (21). Additionally, the estimated size of the 67k calcimedin from the amino acid analysis is 65,000 Da (cysteine residues not determined) as compared to the value of 67,000 Da obtained by PAGE. Comparison of the cDNA derived amino acid compositions of calelectrin and

the lymphocyte protein (Table III) with the 67k calcimedin gave SΔQ values of 187 and 183, respectively. Based on the comparisons between the amino acid analysis or sequence analysis data, the 67k calcimedin appears to be distinct from the lymphocyte protein. This conclusion is supported by the lack of 67k calcimedin antibody staining in lymphocytes (21).

The remaining proteins, with SΔQ values greater than 100, include the liver CAB63 protein (calregulin) (5), the lung 67E protein (22), and calmodulin (23). Each of these proteins is clearly unrelated to any of the other calcium-binding proteins tested. Interestingly, the CAB63 protein showed a correlation coefficient of 48 when compared to rabbit skeletal muscle calsequestrin (24) (data not shown). When compared with a one-to-one heterodimer of calsequestrin (M_r = 45,000) and calmodulin (M_r = 17,000), the CAB63 SΔQ value decreased to 9, suggesting near identity. Addition or subtraction of calmodulin did not substantially alter the SΔQ values for any of the other proteins.

The amino acid data are therefore consistent with the following: 1) a major group sharing homology with the 32,000 dalton lipocortin and calpactin I include Torpedo calelectrin, brain calelectrin, the aorta protein and possibly the lung 67R protein; 2) the 67k calcimedin appears to be distinct from this group of proteins and 3) CAB63, lung 67E and calmodulin are totally unrelated proteins. This analysis demonstrates the utility of the SΔQ method for determining relatedness and ancestry of proteins, especially when amino acid or DNA sequences are not yet available. The groupings of these proteins based on SΔQ values do not necessarily coincide with the groupings defined by less specific criteria such as molecular size, mobility on SDS gel electrophoresis, and binding to immobilized ligands. These other criteria are useful, but not fully adequate, to suggest similarities between proteins and may be misleading if solely relied upon.

Acknowledgements

I thank Dr. S. S. Morse, The Rockefeller University, for critical comment on the manuscript. Supported by The American Heart Association, National Program. The author is an Investigator of the American Heart Association, New York City Affiliate.

References

1. Moore, P.B. (1986) Biochem. J. 238:49-54.
2. Südhof, T.C., Ebbecke, M., Walker, J.H., Fritsche, U., and Boustead, C. (1984) Biochemistry 23:1103-1109.
3. Creutz, C.E., Dowing, L.G., Sando, J.J., Vilar-Plasi, C., Whipple, J.H., and Zaks, W.J. (1983) J. Biol. Chem.258:14664-14674.
4. Owens, R.J., and Crumpton, M.J. (1984) Biochem. J. 219:306-316.
5. Waisman, D.M., Salimath, B.P., and Anderson, M.J. (1985) J. Biol. Chem. 260:1652-1660.
6. Dedman, J.R. (1986) Cell Calcium 7:297-307.
7. Martin, F., Derancourt, J., Capony, J.P., and Cavadore, J.D. (1987) J. Muscle Res. & Cell Motil. 8:77.
8. Geisow, M.J., Walker, J.H., Boustead, C., and Taylor, W. (1987) Bioscience Reports 7:289-307.
9. Südhof, T.C., Slaughter, C.A., Leznicki, I., Barjon, P. and Reynolds, G.A. (1988) Proc. Nat. Acad. Sci. USA 85:664-668.
10. Crompton, M.R., Owens, R.J., Totty, N.F., Moss, S.E., Waterfield, M.D. and Crumpton, M.J. (1988) The EMBO J. 7:21-27.
11. Marchalonis, J.J., and Weltman, J.K. (1971) Comp. Biochem. Physiol. 38B, 609-625.
12. Weltman, J.K., and Dowben, R.M. (1973) Proc. Nat. Acad. Sci. USA 70:3230-3234.
13. Marchalonis, J.J. (1972) Nature New Biol. 236, 84-86.
14. Prager, E.M., and Wilson, A.C. (1971) J. Biol. Chem. 246:5978-5989.

15. Robert, L., Sayolle, J., Derovette, F., and Zabriskie, J. (1972) Transplant. Proc. 4:415-418.
16. Zagalsky, P.S. (1972) Comp. Biochem. Physiol. 41B: 385-397.
17. Gerke, V., and Weber, K. (1984) The EMBO J. 3:227-233.
18. Geisow, M.J., Fritsche, U., Hexham, J.M., Dash, B., and Johnson, T. (1986) Nature 320:636-638.
19. Martin, F., Derancourt, J., Capony, J.P., Colete, S., and Cavadore, J.D. (1987) Biochem. Biophys. Res. Commun. 145:961-968.
20. Rhoads, A.R., Lulla, M., Moore, P.B., and Jackson, C.E. (1985) Biochem. J. 229:587-593.
21. Morse, S.S. and Moore, P.B. (1988) Biochem. J. 251:171-174.
22. Fauvel, J., Vicendo, P., Roques, V., Ragab-Thomas, J., Granier, C., Vilgrain, I., Chambaz, E., Rochet, H., Chap, H., and Douste-Blazy, L. (1987) FEBS Lett. 221:397-402.
23. Dedman, J.R., Potter, J.D., Jackson, R.L., Johnson, J.D., and Means, A.R. (1971) J. Biol. Chem. 252:8415-8422.
24. Slupsky, J.R., Ohnishi, M., Carpenter, M.R., and Reitmeier, R.A.F. (1987) Biochem. 26:6539-6544.
26. Rothhut, B., Comera, C., Prieur, B., Errasfa, M., Minassian, G., and Russo-Marie, F. (1987) FEBS Lett. 219:169-175.