Volume 249, number 1, 8–12 FEB 07153 May 1989

Correlation among the proton charges and molecular masses of myosin subunits

Jean-Jacques Béchet and Anne d'Albis

Laboratoire de Biologie Physicochimique, Unité associée au CNRS 1131, Université de Paris-Sud, 91405 Orsay, France

Received 20 February 1989

The molecular masses and isoelectric points of myosin light and heavy chains were calculated from their known primary sequences and their respective distribution in a two-dimensional graph is displayed. Implications for the electrophoretic study of myosin subunits are inferred from this analysis.

Myosin; Electrophoresis; Molecular mass; Isoelectric point; Light chain; Heavy chain

1. INTRODUCTION

Polyacrylamide gel electrophoresis is a powerful tool for the separation and identification of protein isoforms, and especially of isomyosins [1–3]. These last ones may differ by their light chains and/or heavy chains. Numerous primary sequences of these chains have been elucidated in recent years using chemical and genetic methods [4,5]. The analysis of their molecular masses and of their content in charged amino acid residues may help to understand the electrophoretic behaviour of any newly investigated myosin and that of its subunits, and to predict some of their physicochemical properties.

2. MATERIALS AND METHODS

The molecular masses and the mean ionization constants of amino acids constituting the primary sequences of myosin subunits were found in [6] and [7], respectively. The ionization constant of α-*N*-trimethylalanine was taken to be equal to 10^{-13} [8] while the ionization constants of uncommon amino acids like 3-*N*-methylhistidine, ϵ-*N*-methyllysine, ϵ-*N*-trimethyllysine were assimilated to those of the unsubstituted corresponding amino acids, as a first approximation. The unknown N-termini of some light chains were assumed to be identical to those of light chains from close species while the unknown N-termini of heavy chains were considered to be free. The isoelectric points of the studied proteins or peptides just as their net charges at neutral pH were calculated from the equation discussed by Tanford [9], after determination of the amino acid composition of these molecules.

Correspondence address: J.-J. Béchet, Laboratoire de Biologie Physicochimique, Bâtiment 433, F-91405 Orsay, France

Abbreviations: A1LC, higher molecular mass alkali (or essential) light chain; RLC, regulatory light chain; A2LC, lower molecular mass alkali (or essential) light chain

3. RESULTS AND DISCUSSION

Table 1 sums up the molecular masses, the isoelectric points and the net charges at neutral pH for the essential and regulatory light chains from different species, and table 2 sums up the same features for various myosin heavy chains and their constitutive subfragments. A few experimental values of isoelectric points as reported in the literature are also given; they are usually higher than the calculated values by about 0.2, probably due to the solvent used in these determinations. In fig.1, the distribution of molecular mass and isoelectric point values for the different families of myosin light chains is displayed in a two-dimensional graph reminiscent of the usual two-dimensional electrophoretic pattern [21]. Fig.2 is a comparable graph drawn for myosin heavy chains and their constitutive peptides. It is obvious from

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

Table 1

Calculated molecular masses (MW), isoelectric points (p*I*) and charges (*Z*) at neutral pH for myosin light chains

Origin[a]	A1LC			RLC			A2LC		
	MW	p*I*	*Z*	MW	p*I*	*Z*	MW	p*I*	*Z*
1	20860	5.0 (5.22;	−8	18937	4.81 (4.91)[b]	−8	16569	4.5 (4.72)[b]	−13
2		5.16)[b]		19065	4.99	−6			
3	21086	5.01	−8				16606	4.57	−12
4	20810	5	−8	18750	4.76	−9	16620	4.48	−14
5	20759	5.04 (5.26)[c]	−8	18880	4.81 (5.08)[c]	−8	16524	4.59 (4.73)[c]	−12
6	20505	5.03 (5.5)[d]	−8				16510	4.57 (4.85)[d]	−12
7	21796	5.19 (5.36)[b]	−7	18616	4.85	−9			
8				18700	4.85	−9			
9	21843	5.075	−8						
10	21877	5.01	−9						
11				18791	4.86 (5.09)[c]	−9			
12	21027	5.01	−8						
13	21519	4.97	−9						
14	20487	5.08	−8						
15				19755	4.74	−12	16883	4.42	−15
16							16879	4.52	−13
17							17616	4.26	−20
18				17568	4.53	−11	17603	4.21	−21
19				17249	4.5	−12			
20				17582	4.5	−12			
21				18425	4.61	−10			
22				17662	4.51	−11			
23				17271	4.2	−16	17002	4.4	−17
24				18035	4.4	−12			
25				17599	4.35	−15	18100	4.38	−18
26				23535	4.68	−12	17434	4.22	−20
27							18428	5.17	−6
28							16123	4.39	−12
Mean	21143	5.04	−8	18613	4.64	−11	17064	4.48	−15
± SD	± 528	± 0.06	± 1	± 1460	± 0.22	± 3	± 672	± 0.24	± 4

[a] The origin of myosin light chains (see references in [4,10]) is the fast skeletal muscle of rabbit (1,2), man (3), chicken (4), rat (5) and mouse (6); the ventricular muscle of chicken (7,8), man (9; [11] and 10) and rat (11; [12]); the atrial muscle of mouse (12; [13]) and man (13; [11]); the embryonic muscle of chicken (14); the chicken gizzard muscle (15) and the fibroblast cells of chick embryo (16); the scallop striated muscles from different species (17–20); the smooth muscle of scallop (21,22), ascidian (23), clam foot (24); the squid mantle muscle (25) and the *Drosophila* muscle (26); the lower eukaryotes, *Dictyostelium* (27) and *Physarum* (28)

[b] Experimental values from [14]

[c] Experimental values from [15]

[d] Experimental values from [16]

these figures that almost each light chain or heavy chain may be differentiated on the basis of its molecular mass and/or protonic charge; modern improved electrophoretic separation allows indeed the separation of two proteins differing only by one charge [22] or having molecular masses differing by hardly 1% [23].

With regard to the vertebrate light chains, some relationship is found between the two investigated parameters (fig.1): the higher the molecular mass, the higher the isoelectric point. This relationship is less clear for invertebrate light chains which usually have a lower isoelectric point than the vertebrate light chains of the same essential or regulatory family. On the other hand, particular relative positions of the different A1LC isoforms present in striated muscles from mammals (rat, mouse, rabbit, man) have been recognized [24]; this distribution is however not general for it is not found in guinea pig [25] or chicken (fig.1; inset).

Table 2

Calculated molecular masses (MW), isoelectric points (p*I*) and charges (*Z*) at neutral pH for myosin heavy chains and their fragments

Origin[a]	25 K peptide			50 K peptide			23 K peptide		
	MW	p*I*	*Z*	MW	p*I*	*Z*	MW	p*I*	*Z*
1	23032	7.64	1	49290	6.01	−5	19570[b]	10.01[b]	13[b]
2	23040	6.75	0	49356	6.32	−3	19570[b]	10.01[b]	13[b]
3	23038	7.71	1	49183	5.81	−8	19570[b]	10.23[b]	14[b]
4	23425	5.66	−3	48617	6.22	−6	24010	10.47	23
5	23052	8.06	2	49396	6.01	−5	23469	10.23	19
6	23038	7.43	1	50974	5.63	−8	23297	10.49	21
7							23418	10.61	20
8	23127	6.15	−2	50820	5.79	−10	23512	10.25	18
9	14112	6.19	−2	45347	4.95	−14	21867	10.36	25
10	22981	9.16	6	49531	5.59	−10	23805	10.38	18
11	23063	6.45	−1	47651	6.39	−3	22837	10.23	15
12	23055	5.02	−11	49681	6.02	−5			
13	23504	9.03	6						
Mean	23123[d]	7.1	0	49077	5.89	−7	23277[e]	10.38[e]	20[e]
± SD	± 173	± 1.3	± 4	± 1536	± 0.41	± 3	± 667	± 0.14	± 3

Origin[a]	Head			Rod			Heavy chain		
	MW	p*I*	*Z*	MW	p*I*	*Z*	MW	p*I*	*Z*
1	91855[c]	8.55[c]	8[c]						
2	91929[c]	8.66[c]	10[c]						
3	91754[c]	8.33[c]	7[c]						
4	96017	8.83	15	127716	5.16	−52	223715	5.8	−37
5	95881	8.89	16	126884	5.17	−52	222747	5.81	−36
6	97273	8.91	15	131496	5.1	−57	228751	5.59	−42
7									
8	97423	7.98	6	127579	5.21	−45	224984	5.73	−39
9	81290	8.64	10	45900	10.53	26	127172	9.38	36
10	96280	8.97	14	74670	4.98	−34	170932	5.86	−20
11	93515	8.81	11	150232	5.15	−53	243729	5.54	−41
12									
13									
Mean	96065[d,e]	8.72[e]	12[e]	132781[f]	5.16[f]	−52[f]	228785[f]	5.69[f]	−39[f]
± SD	± 1406	± 0.34	± 4	± 9921	± 0.04	± 4	± 8659	± 0.12	± 3

[a] The origin of myosin heavy chains (see [5] unless indicated) is the skeletal muscle of rabbit (1,2), chicken (3; [17]), rat embryo (4) and chick embryo (5); the embryonic smooth muscle of chicken gizzard (6; [18]); the ventricular muscle of rabbit (7, partial sequence; [19]); the skeletal muscle of soil nematode (8); *Acanthamoeba* type IB (9) and type II (10); *Dictyostelium* (11) and yeast (12; partial sequence); the muscle of *Drosophila* (13, partial sequence; [20])

[b] Only the sequence of 20 K peptide is known. The isoelectric points and net charges of 20 K peptides are usually lower than those of corresponding 23 K peptides by a mean value of about 0.2 and 5, respectively

[c] Partial head constituted by the 25 K, 50 K and 20 K peptides

[d] Values for the fragments of *Acanthamoeba* myosin IB are omitted in the calculation

[e] The three first values in the column are omitted in the calculation

[f] Values for the fragments of *Acanthamoeba* myosins IB and II are omitted in the calculation

These fragments are currently named starting from the N-terminus as the 25 K peptide, the 50 K peptide, the 23 K peptide (including the 20 K peptide) and the rod. The three first fragments as defined from [5] constitute the myosin head

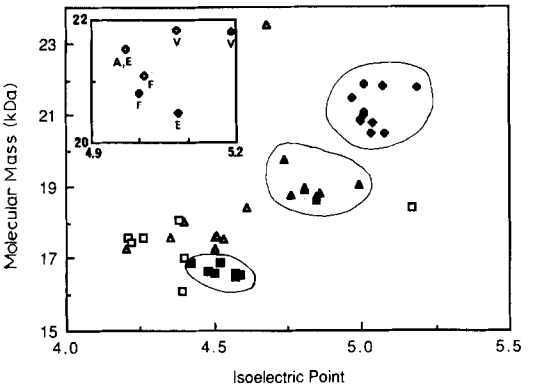

Fig.1. Molecular mass as a function of isoelectric point for myosin light chains from different species (values from table 1). Higher molecular mass alkali light chains from vertebrates (♦); regulatory light chains from vertebrates (▲) and from invertebrates (△); lower molecular mass alkali light chains from vertebrates (■) and from invertebrates (□). The points for some particular family of light chains are surrounded by a faint line. In the inset is the compared distribution of A1LC isoforms present in fast skeletal (F), ventricular (V), atrial (A) and embryonic (E) muscles of chicken (♦) and man (◇).

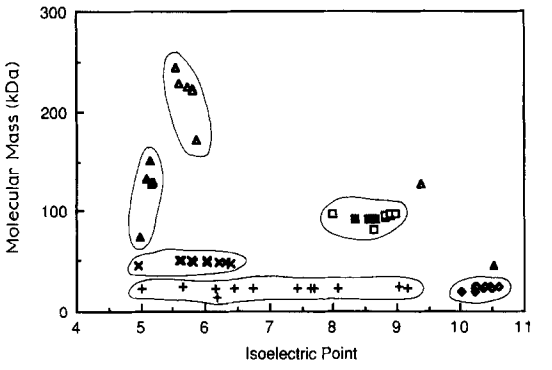

Fig.2. Molecular mass as a function of isoelectric point for myosin heavy chains from different species, and their constitutive fragments (values from table 2). 25 K fragment (+), 50 K fragment (×), 20 K fragment (♦), 23 K fragment (◇), partial head [25 K + 50 K + 20 K] (■), myosin head [25 K + 50 K + 23 K] (□), rod (▲) and myosin heavy chain (△). The points for some particular family of fragments are surrounded by a faint line. Notice the abnormal points for the rod and heavy chain of *Acanthamoeba* myosin IB.

With respect to the isoelectric points of myosin heavy chain subfragments (fig.2), a large dispersion of values around the neutrality is found for the 25 K fragment. This variation, as it was verified, is not linked to the little sequence conservation in the N-terminus of this fragment [5]. Moreover, high values are found for the 20 K (or 23 K) peptide; as suggested by Onishi et al. [26], the basicity of this fragment may play a major role in the affinity of acidic light chains for the heavy chain. At last, the values of the isoelectric points are almost constant in the rod, except for that of the non-filamentous myosin IB from *Acanthamoeba*, while the length of the rod is very variable in the different myosins; this result is linked to the presence of numerous positively or negatively charged residues arranged in alternating bands in the rod [27], and necessary to the stabilisation of myosin filaments.

Thus myosin molecule is characterized by a marked polymorphism which has probably a functional significance. At one and the same time light and heavy chains have variable charges and molecular masses in different myosins; this is responsible for the frequently facilitated separation of close isomyosins in polyacrylamide gel electrophoresis under native conditions.

REFERENCES

[1] Hoh, J.F.Y., McGrath, P.A. and White, R.I. (1976) Biochem. J. 157, 87–95.
[2] D'Albis, A., Pantanoli, C. and Béchet, J.-J. (1979) Eur. J. Biochem. 99, 261–272.
[3] Staron, R.S. and Pette, D. (1987) Biochem. J. 243, 687–699.
[4] Grand, R.J.A. (1982) Life Chem. Rep. 1, 105–160.
[5] Warrick, H.M. and Spudich, J.A. (1987) Annu. Rev. Cell Biol. 3, 379–421.
[6] Mooz, E.D. (1976) in: Handbook of Biochemistry and Molecular Biology – Proteins (Fasman, G.D. ed.) vol.1, pp.111–174, CRC Press, Cleveland.
[7] Fersht, A. (1985) Enzyme Structure and Mechanism, 2nd edn, p.156, W.H. Freeman and Co., New York.
[8] Henry, G.D., Trayer, I.P., Brewer, S. and Levine, B.A. (1985) Eur. J. Biochem. 148, 75–82.
[9] Tanford, C. (1962) Adv. Prot. Chem. 17, 69–165.
[10] Béchet, J.-J. and Houadjeto, M. (1989) Biochim. Biophys. Acta, in press.
[11] Kurabayashi, M., Komuro, I., Tsuchimochi, H., Takaku, F. and Yasaki, Y. (1988) J. Biol. Chem. 263, 13930–13936.
[12] Henderson, S.A., Xu, Y.-C. and Chien, K.R. (1988) Nucleic Acid Res. 16, 4722.
[13] Barton, P.J.R., Robert, B., Cohen, A., Garner, I., Sassoon, D., Weydert, A. and Buckingham, M.E. (1988) J. Biol. Chem. 263, 12669–12676.
[14] Scordilis, S.P. and Mahoney, C.W. (1981) Biophys. J. 33, 277a.
[15] Silver, P.J. and Stull, J.T. (1982) J. Biol. Chem. 257, 6137–6144.
[16] D'Albis, A., Janmot, C. and Béchet, J.-J. (1985) Biochem. Biophys. Res. Commun. 128, 94–100.

[17] Maita, T., Hayashida, M., Tanioka, Y., Komine, Y. and Matsuda, G. (1987) Proc. Natl. Acad. Sci. USA 84, 416–420.
[18] Yanagisawa, M., Hamada, Y., Katsuragawa, Y., Imamura, M., Mikawa, T. and Masaki, T. (1987) J. Mol. Biol. 198, 143–157.
[19] Kavinsky, C.J., Umeda, P.K., Levin, J.E., Sinha, A.M., Nigro, J.M., Jakovcic, S. and Rabinowitz, M. (1984) J. Biol. Chem. 259, 2775–2781.
[20] Wassenberg, D.R., ii, Kronert, W.A., O'Donnell, P.T. and Bernstein, S.I. (1987) J. Biol. Chem. 262, 10741–10747.
[21] O'Farrell, P.H. (1975) J. Biol. Chem. 250, 4007–4021.
[22] Sadano, H., Taniguchi, S., Kakunaga, T. and Baba, T. (1988) J. Biol. Chem. 263, 15868–15871.
[23] Rushbrook, J.I., Wadewitz, A.G., Elzinga, M., Yao, T.-T. and Somes, R.G., jr (1988) Biochemistry 27, 8953–8958.
[24] Barton, P.J.R. and Buckingham, M.E. (1985) Biochem. J. 231, 249–261.
[25] D'Albis, A. and Janmot, C. (1989) Comp. Biochem. Physiol., in press.
[26] Onishi, H., Maita, T., Miyanishi, T., Watanabe, S. and Matsuda, G. (1986) J. Biochem. (Tokyo) 100, 1433–1447.
[27] McLachlan, A.D. (1984) Annu. Rev. Biophys. Bioeng. 13, 167–189.