THE MOLECULAR BASIS OF THE "RED CELL" ACID PHOSPHATASE POLYMORPHISM

J. Dissing

Institute of Forensic Genetics, University of Copenhagen,
Frederik den Femtes Vej 11, DK-2100 Copenhagen, Denmark.

Introduction.

The "classical" genetic markers, poetically referred to as "old friends" by Dr. Sensabaugh, include the fascinating enzyme "red cell" acid phosphatase (ACP1). This polymorphism was discovered in 1963 by Hopkinson et al. [1] who used the powerful new starch gel electrophoretic technique introduced by Smithies [2] on human hemolysates. ACP1 was thus the first genetical enzyme polymorphism to be discovered. To date at least 17 allelic genes have been observed at the ACP1 locus located to a distal region (2p25) on the short arm of chromosome 2. The ACP1*A and ACP1*B alleles prevail in most populations, ACP1*C occurs in polymorphic frequencies in European populations and ACP1*R in certain populations in Southern Africa [3-11].

Acid phosphatase was found to possess several puzzling characteristics. These include genotypic variations of enzyme activity levels [12], activity modulation by purines and pteridines [13,14], phosphotransferase activity [15] and stability [16,17]. Furthermore, the allelic ordering of these variations were found to differ. Thus, ACP1*A < ACP1*B < ACP1*C was found for activity levels, while ACP1*B < ACP1*A < ACP1*C was observed at modulation by pteridines and certain purines (e.g. adenine) or ACP1*C < ACP1*A < ACP1*B for modulation by certain other purines (e.g. hypoxanthine). The order ACP1*A, ACP1*B < ACP1*C (viz. phenotype A, AB, B < AC and BC) was observed with respect to heat stability. Another characteristic property is the expression of each allele as two different isozymes [1], termed "fast" (f) and "slow" (s) in accordance with their electrophoretic mobility [18](Fig. 1). These two isozymes were long believed to be conformational isomers [19,20].

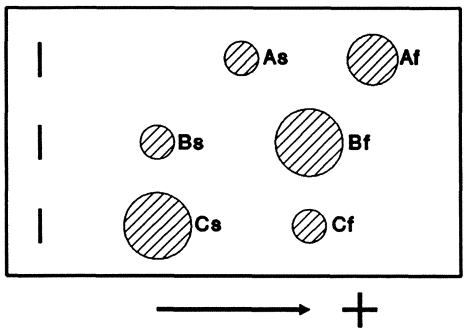


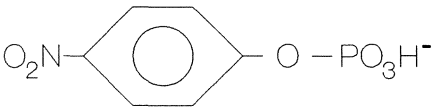
Fig. 1. Electrophoretic patterns of the homozygous ACP1 phenotypes A (top), B and C (citrate buffer, pH 5.9). The allelic **f** and **s** isozymes are denoted **Af** and **As** etc. Application at vertical lines.

Results obtained mainly over the last few years have provided a considerable insight as to the properties and function of the acid phosphatase isozymes and to the structure of the gene. The purpose of the present paper is to review these results and to relate them to the early findings.

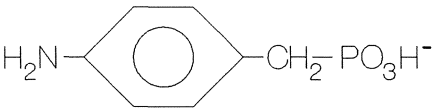
Properties of the isozymes.

Of major importance for the analysis of the acid phosphatase isozymes was the develop-

ment of an efficient purification method by which **f** and **s** isozymes were purified for the first time [21,22]. The procedure involves affinity chromatography on a medium containing the substrate analogue p-aminobenzylphosphonic acid (Fig. 2) and allows the isolation of the acid phosphatase isozymes in mg quantities from red cells with a purification factor of 30,000.



p-nitrophenylphosphate



p-aminobenzylphosphonic acid

Fig. 2. Substrate (top) and substrate analogue of "red cell" acid phosphatase.

Using antibodies produced against homogeneous isozymes it was found that **f** and **s** isozymes of each genetic type contain different as well as common epitopes. No immunochemical differences were observed between the **Af, Bf** and **Cf** isozymes or between the **As, Bs** and **Cs** isozymes [23]. The isozymes were quantitated in red cells from individuals of the various phenotypes using crossed immunoelectrophoresis and it was found that the quantity of enzyme protein paralleled the enzyme activity level, viz. ACP1*A < ACP1*B < ACP1*C ≈ 2:3:4 [23]. Further, it was found that the proportion of **f** and **s** isozyme is allele dependent, the **f:s** ratio being 2:1, 4:1 and 1:4 for the ACP1*A, *B and *C alleles, respectively. The order of this allelic effect (ACP1*C < ACP1*A < ACP1*B) is thus the same as that found at activity modulation. Analysis of the thermostability and the sensitivity to denaturation with urea showed that the **Af, Bf** and **Cf** isozymes have identical stability as have the corresponding three **s** isozymes, whereas the **s** isozymes are significantly more stable than the **f** isozymes [22]. Also in size exclusion chromatography and in SDS polyacrylamide gel electrophoresis differences were observed between **f** and **s** isozymes, whereas identical properties were found within each of these two species of isozymes [22].

A similar pattern was observed with respect to the catalytic properties of the isozymes. The three **f** isozymes (**Af, Bf, Cf**) were indistinguishable in all the measured parameters (specific activity, Michaelis constants, inhibition constants, activity modulation) and the the same was the case for the three **s** isozymes, whereas the **f** and **s** isozymes differed significantly in all the parameters [22,24]. Estimation of the phosphotransferase activity has shown the **Bf** isozyme to be more than twice as active as the **Bs** isozyme [15]. The most striking finding was the diametrically opposite response of **f** and **s** isozymes to modulators; the activity of the **f** isozymes was increased 5.1 fold by hypoxanthine and decreased 40% by adenine, while the activity of the **s** isozymes was unaffected by hypoxanthine but increased 4.6 fold by adenine. Examples of the catalytic properties are given in Table 1. The enzyme kinetic results indicated the presence of a histidine residue in the **s** isozymes, but not in the **f** isozymes, which is of importance for the catalytic function [22].

The important conclusion, therefore, is that the **f** and **s** isozymes differ in all the properties tested, while the **f** isozymes (**Af, Bf, Cf**) are identical in these properties (except electrophoretic mobility) as are the corresponding **s** isozymes, despite the different genetic types. It therefore follows, that from a functional point of view the common phenotypes of acid phosphatase consist of only two different isozymes, **f** and **s**. The allelic differences in quantity and relative proportion of these two isozymes result in the different properties

Table 1. Catalytic parameters for the 6 common acid phosphatase isozymes as determined at pH 6.0, 30^{0}C, using *p*-nitrophenylphosphate as substrate.

				activity modulation (%)[2]	
isozyme	sp.act.[1]	Km (mM)	Ki (mM)[3]	adenine	hypoxanthine
Af	34	0.14	0.47	61	501
Bf	33	0.14	0,44	60	483
Cf	33	0.14	0,43	62	491
As	46	0.53	0,81	383	100
Bs	49	0.51	0,86	407	100
Cs	49	0.52	0,73	407	100

[1]Specific activity is expressed as μmol of *p*-nitrophenol liberated per min per mg of enzyme-protein.
[2]The maximal or minimal activity at saturating concentrations of the modulator expressed as percent of the activity without the modulator.
[3]Inhibition constants of the competitive inhibitor and substrate analogue *p*-aminobenzylphosphonic acid.

of the phenotypes [25]. The genotypic activity variation (ACP1*A < ACP1*B < ACP1*C) is mainly a consequence of the similar variation in the enzyme protein level (but also of the different specific activity of the **f** and **s** isozymes and their relative distribution). The genotypic orders with respect to activity modulation by purines (ACP1*B < ACP1*A < ACP1*C or ACP1*C < ACP1*A < ACP1*B) are accounted for by the different proportions of **f** and **s** isozyme and by the activation or inhibition of these two isozymes exerted by the modulators [24]. Similarly the phenotypic differences in phosphotransferase activity appears to be caused by different properties of **f** and **s** isozymes, as indicated by the higher transferase activity of **f** isozyme relative to **s** isozyme [15]. The higher relative stability of **s** isozymes explains the higher stability of the BC and AC phenotypes as compared to the A, AB and B phenotypes [16,17], since the former phenotypes have a higher content of **s** isozyme [23].

Structure of the isozymes.

A decisive proof against the conformational isomer hypothesis came from unfolding/refolding experiments with the **Bf** and **Bs** isozymes [26]. Unfolding of differently folded isomers of the same polypeptide chain should reduce these to a common molecule, and renaturation of either of the unfolded isomers should result in a mixture of the isomeric forms; however, neither was found to be the case. Peptide mapping revealed that significant structural differences distinguish the two isozymes, but also that large parts of the molecules are identical. These findings were also incompatible with one isozyme being the product of simple posttranslational modifications of the other [26].

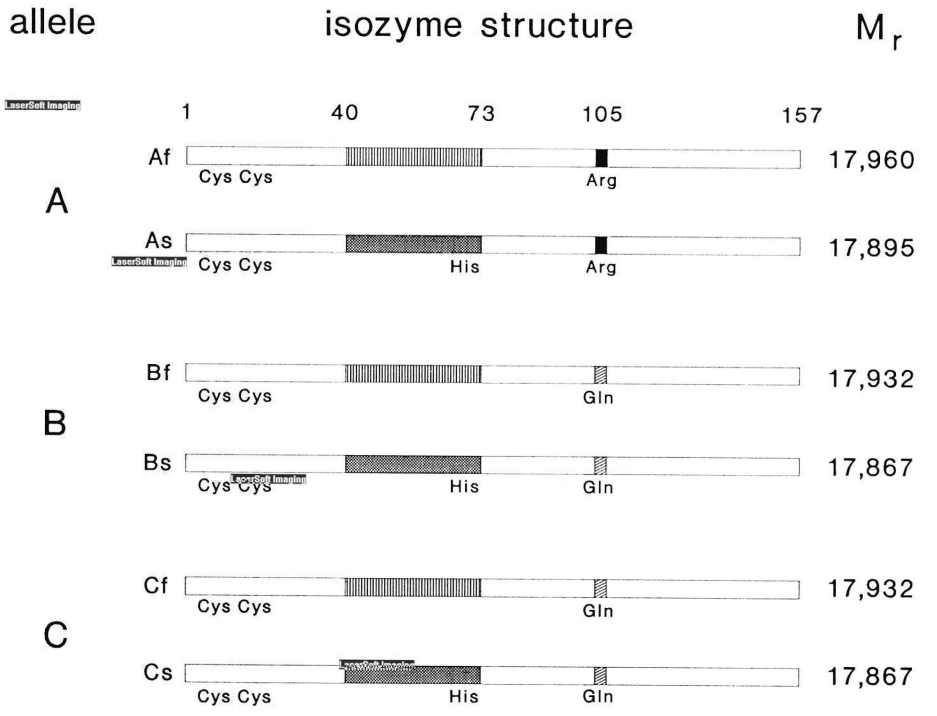


Fig. 3. Diagram of the 6 common isozymes of human "red cell" acid phosphatase.(||||||) **f** specific segment, (xxxxxx) **s** specific segment. Cys-12, Cys-17 and His-69 which all may be at the active site are indicated by Cys and His.

With the determination of the complete amino acid sequence of the **Bf** and **Bs** isozymes and subsequently of the **Af, As, Cf** and **Cs** isozymes a surprising structural relationship between these isozymes was revealed [27-29]. Each isozyme consists of a single non-glycosylated peptide chain of 157 amino acid residues which is acetylated at the amino-terminal alanine. The most striking feature is the existence of a specific **f** or **s** sequence segment, residue 40-73, over which the **f** and **s** isozymes differ substantially (Fig. 3). **f** and **s** isozymes of the same genetic type are identical over the remaining four-fifth of the sequence. Another interesting feature is that the **Bf/Bs** pair of isozymes is identical in sequence with the **Cf/Cs** pair of isozymes. Thus the ACP1*B and ACP1*C alleles encode exactly the same isozymes, the only difference at the protein level being the proportion of **f** and **s** isozyme and the quantity of enzyme protein in the cell [23,29]. The **Af/As** isozyme pair differs from the **Bf/Bs** isozyme pair at a single amino acid residue due to a Glutamine-105 to Arginine substitution. The substitution of a neutral residue with a basic residue in the **A** isozymes explains the higher isoelectric points of the **A** isozymes relative to those of the corresponding **B** and **C** isozymes [30]. Both **f** and **s** isozymes contain 8 cysteine residues at identical positions, all occurring as free sulphydryls. Two of the cysteines appear to be located at the active site [31] (Fig. 3).

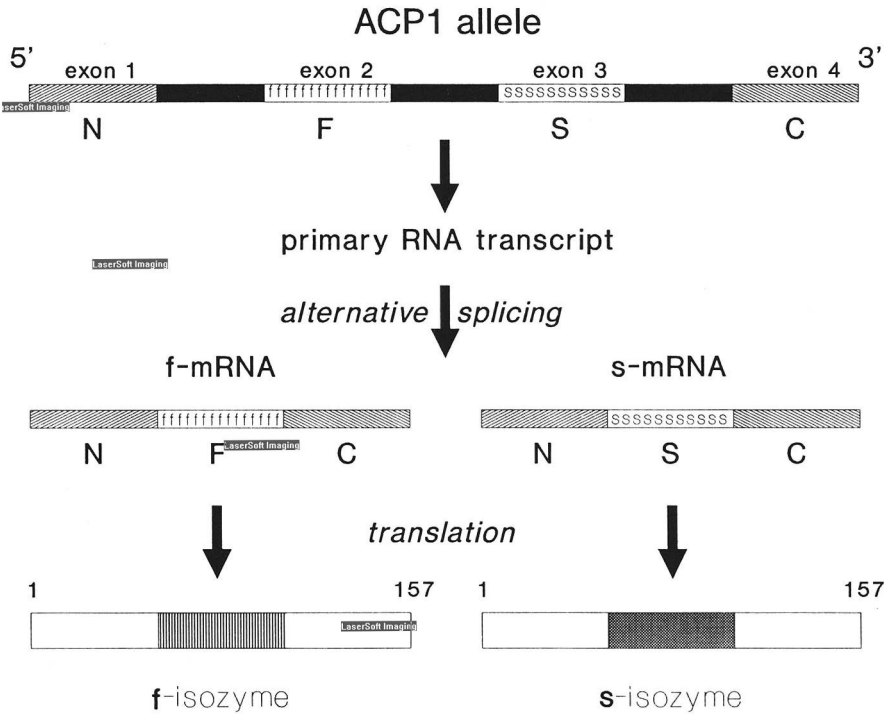


Fig. 4. Proposed mechanism for the generation of **f** and **s** isozymes by ACP1 alleles.

The specific **f** and **s** sequence segments must account for the differences in catalytic and molecular properties between the **f** and **s** isozymes. The **s** specific sequence contains an extra basic residue (His-69) which explains the lesser anodal mobility of the **s** isozymes compared to the corresponding **f** isozymes (Fig. 3). His-69 may also be identical to the histidine residue which appears to participate in the catalytic activity of the **s** isozymes [22]. The **f** specific sequence exhibits a lesser capacity for hydrophobic and a higher capacity for hydrophillic interaction than the **s** specific sequence [28]. This may account for the lower stability of the **f** isozymes relative to the **s** isozymes.

Structure of the ACP1 locus:

Based on the results obtained by peptide mapping of the **Bf** and **Bs** isozymes it has been hypothesized that the generation of **f** and **s** isozymes are the result of alternative splicing of the primary RNA transcript [26]. This hypothesis was further supported by the subsequent determination of the primary structure of the 6 common **f** and **s** isozymes [27-29]. The hypothesis predicted that the ACP1 locus is composed of at least 4 coding sequences, exons (A, F, S, C), interspaced with non-coding sequences, introns. The exons code for the amino-terminal region, the **f** specific region, the **s** specific region and the carboxy-terminal region of the isozymes, respectively. By alternative splicing of the primary RNA transcript two different mRNA molecules, a **f**-mRNA and a **s**-mRNA, are

generated from each allele, the messengers consisting of the N,F,C and the N,S,C exons, respectively (Fig. 4) [26,28]. Mutational events in the N or C exons common to **f**-mRNA and **s**-mRNA would alter the **f** and **s** isozymes to the same extent, precisely as is seen with the Gln-105 to Arg substitution in the **A** isozymes. This offers a simple mechanism for the pairwise variation in the electrophoretic mobility of the **f** and **s** isozymes as observed for the various ACP1 alleles.

Although the splicing model has been questioned by Wo et al. [32] it is strongly supported by recent DNA-sequencing results [33]. Genomic sequence spanning approximately 6 kb of the ACP1 locus was determined and it was found that the **f** and **s** specific sequences are indeed encoded by two distinct exons (E3F and E3S) interspaced by a 41 bp non-coding segment. The E3F and E3S exons are flanked on either side by linearly arranged exons (E2, E4, E5, E6) encoding the amino- and carboxy-terminal sequences identical in the **f** and **s** isozymes (Fig. 5). The six exons account for all the coding sequence except that for amino acid residues 1-13. Genomic sequence 3' to E2 as well as the interior portions of the 3Kb and 1.8Kb introns have yet to be determined. Interestingly the 41 bp intervening sequence between exons 3F and 3S is too short to act as a functional intron [34]; this is consistent with a mechanistic model for mutually exclusive spliced exons in which a transcript segment containing either E3F-I3F or I3F-E3S is spliced out in the formation of the

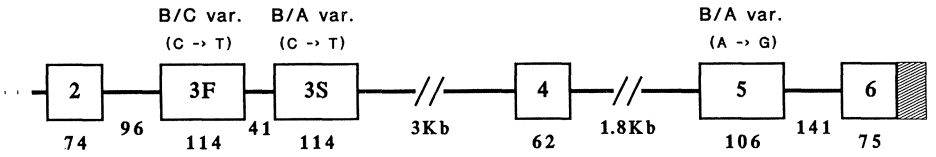


Fig. 5. Scheme of the ACP1 gene locus of human low-molecular-weight acid phosphatase. Boxes represent exons, lines represent introns, and the hatched box represents 3' non-coding sequence. Exon and intron sizes (bp) are indicated below. Allelic base transitions are indicated above the respective exons.

messengers for the **s** and **f** isozymes, respectively. Only three base substitutions, all located internally in exons, have so far been found to distinguish the ACP1*A, ACP1*B and ACP1*C alleles. An A-G transition in codon 105 of E5 accounts for the Gln-105 to Arg substitution that distinguishes the A isozymes from the B and C isozymes. The other two substitutions, both silent C-T transitions, are interesting in that they are located 15 and 12 bases into the 3F and 3S exons, respectively. The former distinguishes the ACP1*C alleles from the ACP1*A and ACP1*B alleles and the latter the ACP1*A alleles from the ACP1*B and ACP1*C alleles. Exon based substitutions located in the vicinity of a splice site may influence splice site recognition [35]. They may therefore provide the basis for allele specific splicing pathways, which result in the quantitative differences of **f** and **s** isozymes as expressed by the ACP1*A, ACP1*B and ACP1*C alleles [23].

DNA-based genotyping:
The allelic sequence differences generate restriction sites that can be used for PCR based genotyping as recently described [36]. However, the sequence differences can also be detected conveniently using the single-strand conformation polymorphism technique introduced by Orita [37]. This is based on the observation that single base substitutions when positioned within a relatively short DNA fragment (e.g. 100-300 bp PCR products) in many cases leads to a change in the electrophoretic mobility of the separated native strands. The reason for this is probably a change in the tri-dimensional conformation of the refolded single strands as caused be the base substitution. Using this technique on 240 bp and 190 bp genomic PCR products encompassing the mutational sites in exons 3F and 3S, respectively, it is possible to determine the common 6 ACP1 phenotypes within a few hours (Fig. 6).

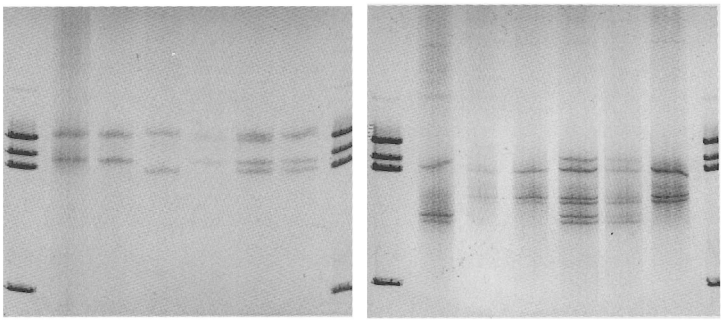


Fig. 6. Determination of classical ACP1 genotypes using PCR and SSCP. The electrophoretic separation was performed in precast 20% polyacrylamide gels (Pharmacia) for 300 Volt hours at 10^0C using the PhastSystem (Pharmacia). Anode at bottom. Lane 1,8,9, 16: Φx174 HAEIII ds size markers (603, 872, 1,078 and 1,353 bp); Lanes 2-7: 240 base ss-DNA encompassing the C-specific mutation at codon 43 in the 3F exon amplified from genomic DNA from individuals of ACP1 phenotype A, B, C, AB, AC and BC, respectively. Lanes 9-15: 190 ss-DNA encompassing the A-specific mutation at codon 41 in the 3S exon amplified from genomic DNA from the same individuals as in lanes 2-7.

Occurrence and physiological importance:
The "red cell" acid phosphatase has been detected in all tissues tested [38,39] and is identical to the cytosolic low-molecular-weight acid phosphatase which has been isolated from diverse vertebrate species [22,40-45]. Thus a high degree of sequence homology has been found between the human, bovine and rat enzymes [28,44,46]. The generation of two isoforms appears to be a general property of low-molecular-weight acid phosphatase genes [29 and references cited therein].

The physiological function (or functions) of the low-molecular-acid phosphatase has yet to be determined. However, certain phosphotyrosyl proteins are hydrolysed by the enzyme and recent results suggest it to play a role in the important control of protein phosphotyrosine levels, which affects cellular processes such as proliferation, differentiation and transformation [47-51]. On the other hand, flavin mononucleotide (FMN) is also a naturally occurring substrate [14,52,53]. By influencing the level of FMN and thereby FAD the low-molecular-weight acid phosphatase may thus affect the activity of flavoenzymes, which are of importance for a number of cellular processes. Given the different properties of **f** and **s** isozymes it is likely that they serve different cellular functions.

Further research will be needed to disclose the precise biological function and a detailed analysis of the ACP1 locus combined with deliberate manipulations of the sequence are required to determine the mechanisms that control the splicing pathway.

References:

1 Hopkinson, D.A., Spencer, N. and Harris, H. (1963) Nature 199, 969-971.
2 Smithies, O. (1955) Biochem. J. 61, 629-641.
3 Povey, S. and Falk, C.T. (1989) Cytogenet. Cell Genet.51, 91-105.
4 Herbich, J., Fisher, R.A. and Hopkinson, D.A. (1970) Ann. Hum. Genet.34, 145-151.
5 Jenkins, T. and Corfield, V. (1972) Ann. Hum. Genet.35, 379-391.
6 Walter, H. (1976) Hum. Hered. 26, 25-33.
7 Santachiara-Benerecetti, A.S., Ranzani, G.N.and Antonini, G.(1977) Am. J. Hum. Genet.29, 635-638.
8 Miller, S.A., Nelson, M.S. and Dykes, D.D. (1987) Hum. Hered. 37, 371-375.
9 Arnaud, J., Vavrusa, B., Sevin, J. and Constans, J. (1989) Hum. Hered. 39, 288-293.
10 Farabegoli, A. and Barbujani, G.(1990) Hum. Hered. 40, 313-321.
11 Roychoudhury, A.K. and Nei, M. (1988) Human polymorphic genes world distribution, Oxford University Press, pp. 109-110.
12 Spencer, N., Hopkinson, D.A. and Harris, H. (1964) Nature 201, 299-300.
13 Sensabaugh, G.F.and Golden,V.L. (1978) Am. J. Hum. Genet.30, 553-560.
14 Mansfield, E. and Sensabaugh, G.F.(1978) in The Red Cell (Brewer, G.J.,ed.), pp. 233-247, Allan R. Liss, New York.
15 Golden,V.L. and Sensabaugh, G.F.(1986) Hum. Genet.72, 340-343.
16 Luffman, J.E. and Harris, H. (1967) Am. J. Hum. Genet.30, 387-400.
17 Shinoda, T. (1968) J. Biochem. 61, 733-741.
18 Dissing, J. (1986) in Advances in Forensic Haemogenetics 1 (Brinkmann, B. and Henningsen, K., eds.), pp. 127-131, Springer, Berlin.
19 Fisher, R.A. and Harris, H. (1971) Ann. Hum. Genet.34, 439-448.
20 Rogers, P.A., Fisher, R.A. and Putt, W. (1978) Biochem. Genet.7, 727-738.
21 Dissing, J., Dahl, O. and Svensmark, O. (1979) Biochim. Biophys. Acta 569, 159-176.
22 Dissing, J. and Svensmark, O. (1990) Biochim. Biophys. Acta 1041, 232-242.
23 Dissing, J. (1987) Biochem. Genet.25, 901-917.
24 Dissing, J., Rangaard, B. and Christensen, U. (1993) Biochim. Biophys. Acta. 1162, 275-282.
25 Dissing, J. (1992) in Advances in Forensic Haemogenetics 4 (Rittner, C. and Schneider, P.M., eds.), pp. 303-305, Springer, Berlin.
26 Dissing, J. and Sensabaugh, G.F.(1987) Biochem. Genet.25, 919-927.
27 Dissing, J. and Johnsen, A.H. (1990) in Advances in Forensic Haemogenetics 3 (Polesky, H.F. and Mayr, W.R., eds.), pp. 211-213, Springer, Berlin.
28 Dissing, J., Johnsen, A.H. and Sensabaugh, G.F.(1991) J. Biol. Chem. 266, 20619-20625.
29 Dissing, J. and Johnsen, A.H. (1992) Biochim. Biophys. Acta 1121, 261-268.
30 Divall, G.B.(1981) Forens. Sci. Int. 18, 67-78.
31 Cirri, P., Chiarugi, P., Camici, G.,Manao, G.,Raugei, G.,Cappugi, G.and Ramponi, G.(1993) Eur. J. Biochem. 214, 647-657.
32 Wo, Y.Y.P., Mccormack, A.L., Shabanowitz, J., Hunt, D.F., Davis, J.P., Mitchell, G.L.and Van Etten, R.L. (1992) J. Biol. Chem. 267, 10856-10865.
33 Ayer Lazaruk, K.D., Dissing, J. and Sensabaugh, G.F.(1993) Biochem. Biophys. Res. Commun. (in press)
34 Ruskin, B., Greene,J.M. and Green,M.R. (1985) Cell 41, 833-844.
35 Cooper, Thomas A. (1992) J. Biol. Chem. 267, 5330-5338.
36 Sensabaugh, G.F.and Lazaruk, K.Ayer (1993) Hum. Mol. Genet.2, 1079.
37 Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. (1989) Genomics5, 874-879.
38 Blake, N.M., Kirk, R.L., Barnes, K.R. and Thompson, J.M. (1973) Jpn. J. Hum. Genet.18, 10-23.
39 Swallow, D.M., Povey, S. and Harris, H. (1973) Ann. Hum. Genet.37, 31-38.
40 Baxter, J.H. and Suelter, C.H. (1985) Arch. Biochem. Biophys. 239, 29-37.
41 Fujimoto, S., Murakami, K., Ishikawa, A., Himi, K. and Ohara, A. (1988) Chem. Pharm. Bull. Tokyo 36, 3020-3026.
42 Zhang, Z.Y. and Van Etten, R.L. (1990) Arch. Biochem. Biophys. 282, 39-49.
43 Saeed, A., Tremori, E., Manao, G.,Camici, G.,Cappugi, G.and Ramponi, G.(1990) Physiol. Chem. Phys. & Med. NMR 22,
44 Manao, G.,Pazzagli, L., Cirri, P., Caselli, A., Camici, G.,Cappugi, G.,Saeed, A. and Ramponi, G. (1992) J. Protein. Chem. 11, 333-345.
45 Panara, F., Angiolillo, A., Secca, T., Dirosa, I., Fagotti, A. and Pascolini, R. (1991) Int. J. Biochem. 23, 1115-1122.
46 Camici, G.,Manao, G.,Cappugi, G.,Modesti, A., Stefani, M. and Ramponi, G.(1989) J. Biol. Chem. 264, 2560-2567.
47 Chernoff, J. and Li, H.-C. (1985) Arch. Biochem. Biophys. 240, 135-145.
48 Boivin, P. and Galand,C. (1986) Biochem. Biophys. Res. Commun. 134, 557-564.
49 Ramponi, G., Manao, G., Camici, G., Cappugi, G., Ruggiero, M. and Bottaro, D.P. (1989) FEBS Letters 250, 469-473.
50 Lau, K.-H.W., Farley, J.R. and Baylink, D.J. (1989) Biochem. J. 257, 23-36.
51 Ruggiero, M., Pazzagli, C., Rigacci, S., Magnelli, L., Raugei, G.,Berti, A., Chiarugi, V.P., Pierce, J.H., Camici, G.and Ramponi, G.(1993) FEBS Lett. 326, 294-298.
52 Tsuboi, K.K. and Hudson, P.B. (1956) Arch. Biochem. Biophys. 61, 197-210.
53 Fuchs, K.R., Shekels, L.L. and Bernlohr, D.A. (1992) Biochem. Biophys. Res. Commun. 189, 1598-1605.