Use of Principal Component Analysis To Study the Relationship between Physical/Chemical Properties and the Milk-Clotting to Proteolytic Activity Ratio of Some Aspartyl Proteinases

Rickey Y. Yada*[1] and Shuryo Nakai

The relationship between physical/chemical properties and the enzymatic activity (milk-clotting to proteolytic activity ratio) of aspartyl proteinases was investigated by using the multivariate statistical technique known as principal component analysis. Principal component analysis of various structural and intrinsic properties of aspartyl and non-aspartyl proteinases resulted in the formation of distinct groups. When the parameters that contribute to the principal components were examined, it was possible to conclude that secondary structural parameters had a pronounced effect on classification. Examination of the milk-clotting to proteolytic activity ratios, in relation to the classifications, generally indicated that proteinases with similar activities were grouped closer together.

INTRODUCTION

The proteinases that are successively used in the cheese-making process belong to the class of enzymes known as aspartyl proteinases. These enzymes are characterized by the presence of two aspartic acid residues in the active site as well as having a general optimal activity in the pH range of 1.5–5.0, depending on the particular enzyme and substrate combination being studied (Voynick and Fruton, 1971; Dalgleish, 1982). In addition, sequence and structural homology have been demonstrated for a number of the aspartyl proteinases (Foltmann and Pedersen, 1976; Hsu et al., 1977).

The success of an enzyme used in cheese making lies not only in its ability to clot milk but in the relationship between milk-clotting ability and general proteolytic ability (Dalgleish, 1982). Even within the aspartyl proteinase class, the enzymes show dissimilarity in their action. Chymosin has demonstrated a high milk-clotting to proteolytic activity ratio, whereas porcine pepsin has a low ratio. This difference in activity may be the result of subtle conformational differences (Visser, 1981). The ability to quantitate relationships between structure and function of a molecule is a problem that has continuously perplexed scientists. It is generally accepted that the function of a protein cannot be understood until its structure is known (Schulz et al., 1974). Hydrophobic, electrostatic, and steric forces are three parameters that affect structure and may be used to predict function (Stuper et al., 1979). Although previous researchers have measured various physical and structural properties of some aspartyl proteinases, no attempts have been made to quantitate the relationship between structure and function. The objectives of the present paper were as follows: first, to measure various structural and intrinsic properties of some aspartyl as well as some non-aspartyl proteinases; second, to attempt to classify the proteinases based on these properties using the multivariate technique known as principal component analysis and relate the classifications obtained to the milk-clotting to proteolytic activity ratio of the various proteinases.

MATERIALS AND METHODS

Porcine pepsin (pepsin A, EC 3.4.23.1) two times crystallized and chymosin (rennin, EC 3.4.23.4) were obtained from Sigma Chemical Co. (St. Louis, MO). *Mucor miehei* proteinase and *Endothia parasitica* proteinase were courtesy of Pfizer Canada (Kirkland, PQ). *Mucor pusillus* var. Lindt proteinase was a generous gift from Dr. S. Iwasaki (Meito Sangyo Co., Ltd., Tokyo, Japan). *Aspergillus saitoi* proteinase was obtained from Calbiochem (San Diego, CA). Crystals of penicillopepsin were supplied by Dr. T. Hofmann (University of Toronto, Toronto, ON). The non-aspartyl proteinases, namely papain (two times crystallized), trypsin (two times crystallized from bovine pancreas), and α-chymotrypsin (three times crystallized from bovine pancreas) were obtained from Sigma Chemical Co. (St. Louis, MO). The purity of the proteinases was determined on the basis of the densitometric scans of SDS polyacrylamide electrophoretic gels (Laemmli, 1970). From this criteria the following purities were obtained: pepsin, 90%; chymosin, 91%; *M. miehei* proteinase, 89%; *M. pusillus* proteinase, 90%; *E. parasitica* proteinase, 86%; *A. saitoi*, 81%; penicillopepsin, 98%; trypsin, 89%; α-chymotrypsin, 92%; papain, 87%.

Circular Dichroism. CD spectra were measured on a Jasco J-500A spectropolarimeter (Japan Spectronic Co., Ltd., Tokyo, Japan) under a constant nitrogen flush at 20 °C. The instrument was calibrated by a two-point calibration technique at wavelengths 290.5 and 192.5 nm using a 600 mg/L solution of camphor-d_{10}-sulfonic acid (Chen and Yang, 1977).

1. Sample Preparation. Proteinases were dissolved in either phosphate (sodium phosphate) or acetate buffer (sodium acetate–acetic acid) ($\gamma/2 = 0.01$) depending on the pH of the assay. Acetate buffer was used for pH values 5.0, 5.3, and 5.8, while phosphate buffer was used for pH values 6.3, 7.0, and 8.0. The protein solutions were then filtered through a 0.45-μm Millex-HA filter (Millipore Corp., Bedford, MA) prior to CD analysis. Analyses were carried out within 1.0 h of sample preparation.

2. Far-UV Spectra (190–240 nm). Between 190 and 240 nm, a 1.0-mm cell was used with a protein concentration of approximately 0.1 mg/mL in the appropriate buffer. The spectral data were reported in terms of [Θ]; the molar ellipticity per residue was calculated from eq 1, where λ

$$[\Theta]_{MRW\lambda} = \frac{\Theta MRW}{10dc} \quad (1)$$

= wavelength, Θ = observed ellipticity (deg), MRW = mean residue weight, c = concentration (g/mL), and d = path length (cm). The following mean residue weights were used: 96, chymosin (Foltmann, 1966, 1981); 107, pepsin (Fruton, 1970; Foltmann, 1981); 103, *M. miehei*

Department of Food Science, University of British Columbia, Vancouver, British Columbia, Canada V6T 2A2.
[1]Present address: Department of Food Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1.

0021-8561/86/1434-0675$01.50/0

proteinase (Ottesen and Rickert, 1970; Bech and Foltmann, 1981); 109, *M. pusillus* (Arima et al., 1970); 114, *E. parasitica* proteinase (Whitaker, 1970); 119, *A. saitoi* (Ichishima and Yoshida, 1966; Ichishima, 1970); 99, penicillopepsin (Sodek and Hofmann, 1970; Hsu et al., 1977); 108, trypsin (Cunningham et al., 1953; Walsh and Neurath, 1964); 98, α-chymotrypsin (Sober, 1970; Chang, et al., 1978); 98, papain (Smith and Kimmel, 1960; Chang et al., 1978). The molar ellipticity values were not corrected for the refractive index factor.

Average Hydrophobicities. Average hydrophobicities [H_ϕ(av)] for the various proteinases were calculated based on the algorithm of Bigelow (1967).

Hydrophobicity Using Fluorescent Probes. Protein hydrophobicity was determined by using the fluorescent probes *cis*-parinaric acid (CPA) and the magnesium salt of 1-anilino-8-naphthalenesulfonic acid (ANS). For the determination of hydrophobicity using CPA, the method of Townsend and Nakai (1984) was employed while the method of Hayakawa and Nakai (1985) was employed for the determination of hydrophobicity using ANS. Duplicate determinations were made for each probe.

Acidic to Basic Acid Ratio. On the basis of amino acid composition data, the acidic (Asp, Glu) to basic (Lys, His, Arg) amino acid ratio was calculated.

ζ Potential. The ζ potential measurements of the various proteinases at different pH values were measured according to the method of Hayakawa and Nakai (1985).

Accessible Surface Area. The accessible surface area (ASA) of each proteinase was calculated on the basis of their molecular weights according to the algorithm of Janin (1976).

Determination of Milk-Clotting and Proteolytic Activities. The milk-clotting activity of the various proteinases over the pH range (pH 5.0–8.0) was determined according to the method of Iwasaki et al. (1967). Berridge (1945) defined 1 unit of milk-clotting activity as the amount of proteinase that clots 10 mL of reconstituted skim milk in 100 s at 30 °C. The specific activity was expressed as milk-clotting activity per milligram of protein. In order to activate papain for milk-clotting measurement, the proteinase solution was prepared in the presence of 0.05 M cysteine and 0.02 M EDTA (Arnon, 1970).

The proteolytic activity of the various aspartyl proteinases to hydrolyze sodium caseinate over the pH range of 5.0–8.0 was determined by using a method described by Green (1972) with slight modifications. The substrate solution contained 1% (w/v) sodium caseinate in either 0.1 M acetate or 0.1 M phosphate buffer depending on the pH. To 2.5 mL of substrate, preincubated at 35 °C for 10 min, was added 0.5 mL of enzyme and the mixture incubated at 35 °C for 10 min; 2.5 mL of 6.6% (w/v) trichloroacetic acid (TCA) was then added to give a final TCA concentration of 3%. The resulting precipitate was removed by filtration through Whatman No. 2 paper. The filtrate was then analyzed for free amino groups according to the method of Kwan et al. (1983) with slight modifications. To 0.1 mL of filtrate was added 0.3 mL of 1 M K_2HPO_4 first, followed by 0.15 mL of 0.03% (w/v) fluorescamine in acetone, which was added rapidly and directly to the solution and mixed immediately on a Vortex mixer. A volume of 3.0 mL of distilled–deionized water was added to the reaction mixture and the mixture vortexed once again. The fluorescence of the final mixture was measured with an Aminco Bowman 4-8202 spectrofluorometer using an excitation wavelength of 395 nm and an emission wavelength of 480 nm. Blanks were prepared following the same procedure except that TCA was added to the substrate prior to the addition of the enzyme solution. A standard curve was constructed against tyrosine. Proteolytic activity was defined as the amount of tyrosine released per milligram of enzyme. Papain was activated as previously described for milk-clotting activity.

Principal Component Analysis. Data collected for the various proteinases were subjected to principal component analysis using the BMDP:4M program (Frane et al., 1981) and run on an Amdahl 470 V/8 computer. Sixteen variables were included: molar ellipticity values at 10 wavelengths from the CD spectra (190, 193, 198, 200, 202, 210, 213, 222, 224, 225 nm), ζ potential, acidic/basic amino acid ratio, Bigelow average hydrophobicity, accessible surface area, CPA hydrophobicity, ANS hydrophobicity. The 10 wavelengths were selected on the basis of characteristic wavelengths representative of each of the secondary structure fractions (Chang et al., 1978).

RESULTS AND DISCUSSION

Data collected for the various proteinases examined in the present study (Table I) were subjected to principal component analysis (PCA) in attempts to identify parameters important for the classification of these proteinases.

When the 16 original variables for each of the proteinase samples were entered in the principal component analysis program, four factors were returned that met the criteria of their eigenvalues exceeding 1.0 (Table II). Aishima (1979a,b,c) used the same criteria for selecting principal components. The sum of the squares of the factor loadings computed for the original variables (Table III) is equal to the eigenvalue of each factor (Table II), thus representing the extent of contribution of each factor to the total variance. It is customary to consider only those factors that have eigenvalues of 1.0 or greater as having any practical significance (Jeffers, 1967). The four factors or principal components obtained accounted for more than 85% of the total variance (Table II), where the total variance is the sum of the individual variances for each of the original variables. The factor loadings for each factor give an indication of the importance of the original variables to that factor. The factor loadings are presented in Table III. Examination of the factor loadings indicated that factor 1 was primarily concerned with circular dichroism spectral data. Variables contributing to factor 2 included molar ellipticity values at wavelengths 193, 198, 200, and 202 nm. It is interesting to note that these wavelengths correspond to a characteristic wavelength for each of the major secondary structure fractions. The CD band at 193 nm is one of the bands characteristic of α-helix, the 198-nm band characteristic of the β-sheet, the 200-nm band characteristic of the unordered fraction, and the 202-nm band characteristic of the β-turn fraction. Factor 3 was characterized by ANS hydrophobicity, acidic to basic amino acid ratio, ζ potential, and CPA hydrophobicity. Variables contributing to factor 4 included Bigelow average hydrophobicity, ANS hydrophobicity, and accessible surface area, as well as CPA hydrophobicity.

Two-dimensional plots involving the first three factors are presented in Figures 1 and 2. The plot of factor 2 vs. factor 1 (Figure 1) indicated that the proteinases from microbial sources (i.e., *M. miehei*, *M. pusillus*, *E. parasitica*, and *A. saitoi* proteinases and pencillopepsin) were closely associated with one another, while pepsin at pH values 5.0–6.3 formed one group and chymosin at pH values 5.8, 6.3, and 7.0 formed another group. The non-aspartyl proteinases (i.e., papain, α-chymotrypsin, and trypsin) were not associated with any group. In comparison to the aspartyl proteinases examined, the non-aspartyl proteinases had extremely low milk-clotting to proteolytic

Table I. Physical/Chemical Measurements of Several Proteolytic Enzymes

		$[\Theta]_{MRW\lambda}$ ($\times 10^{-3}$), deg cm^2 dmol^{-1}																
protein	pH	190	193	198	200	202	210	213	222	224	225[a]	ZP[b]	H_ϕ(av)[c]	CHG[d]	ASA[e]	ANS[f]	CPA[g]	MC/PA[h]
M. miehei	5.0	97	−63	182	−159	−1040	−2711	−2787	−1398	−1103	−953	−14.7	1109	2.0	13000	2.0	21.0	114.31
proteinase	5.3	232	408	130	−443	−999	−2589	−2574	−1456	−958	−833	−19.2	1109	2.0	13000	2.0	21.0	101.96
	5.8	428	177	489	−248	−1120	−2690	−2730	−1418	−1035	−833	−26.0	1109	2.0	13000	2.0	21.0	97.58
	6.3	313	−6	62	−356	−1286	−2811	−2820	−1575	−1168	−952	−32.9	1109	2.0	13000	2.0	21.0	71.05
	7.0	−762	−21	−126	−602	−1320	−3003	−2962	−1736	−1206	−915	−38.4	1109	2.0	13000	2.0	21.0	6.93
	8.0	643	−51	−364	−498	−1261	−2717	−2725	−1622	−1263	−1020	−45.3	1109	2.0	13000	2.0	21.0	0.00
E. parasitica	5.0	1400	1402	542	86	−734	−2363	−2308	−1442	−1151	−1031	−9.3	923	2.4	12900	7.0	113.0	82.11
proteinase	5.3	866	1198	339	−31	−636	−2163	−2236	−1322	−1094	−920	−14.9	923	2.4	12900	7.0	113.0	62.52
	5.8	1563	1303	683	−204	−731	−2203	−2423	−1542	−1272	−1079	−20.5	923	2.4	12900	7.0	113.0	62.49
	6.3	1141	1414	385	26	−836	−2492	−2377	−1462	−1193	−1050	−26.1	923	2.4	12900	7.0	113.0	60.36
	7.0	192	383	−4	−496	−938	−2237	−2250	−1343	−1061	−952	−30.1	923	2.4	12900	7.0	113.0	0.00
	8.0	48	436	−697	−1486	−1889	−2537	−2492	−2018	−1790	−1670	−33.9	923	2.4	12900	7.0	113.0	0.00
chymosin	5.0	1667	1337	−1449	−3373	−5133	−7438	−7015	−5589	−4813	−4532	−6.9	1120	1.6	11300	48.0	96.0	159.61
	5.3	1978	1604	−3359	−4897	−5745	−6434	−5798	−4610	−4209	−3961	−9.5	1120	1.6	11300	48.0	96.0	132.81
	5.8	7320	8960	2692	−906	−3971	−9761	−9391	−8191	−7485	−7139	−11.7	1120	1.6	11300	48.0	96.0	97.06
	6.3	8912	8913	2788	−1125	−4610	−10460	−9577	−8393	−7952	−7656	−14.0	1120	1.6	11300	48.0	96.0	95.75
	7.0	9650	8655	2454	−1308	−4149	−9680	−9399	−8135	−7444	−7037	−17.4	1120	1.6	11300	48.0	96.0	13.64
	8.0	8171	6488	406	−3104	−5853	−10048	−9476	−8115	−7100	−6793	−20.8	1120	1.6	11300	48.0	96.0	0.00
pepsin	5.0	4726	5028	2591	494	1935	−7245	−7357	−5700	−4158	−3756	−20.2	1063	10.7	12300	1.0	6.0	74.43
	5.3	3414	4514	2423	29	−2149	−7182	−7178	−4768	−3892	−3517	−25.0	1063	10.7	12300	1.0	6.0	57.89
	5.8	3392	4174	1963	21	−2265	−7014	−6702	−4795	−3857	−3465	−29.5	1063	10.7	12300	1.0	6.0	53.12
	6.3	2684	4094	2051	−47	−2398	−7144	−7059	−4953	−3893	−3402	−34.6	1063	10.7	12300	1.0	6.0	47.43
	7.0	−436	−186	−5199	−6414	−7350	−6770	−6042	−4102	−3444	−3136	−41.0	1063	10.7	12300	1.0	6.0	0.00
	8.0	−1776	−2295	−5842	−7088	−7528	−6847	−6086	−4038	−3511	−3255	−52.2	1063	10.7	12300	1.0	6.0	0.00
M. pusillus	5.0	19	−869	−366	−1014	−1842	−3030	−3060	−1477	−1063	−820	−7.47	1041	3.6	30600	7.0	3.0	127.30
proteinase	5.3	1228	−522	−879	−1323	−2029	−3000	−2918	−1434	−942	−712	−9.91	1041	3.6	30600	7.0	3.0	90.01
	5.8	532	−237	−429	−1184	−1884	−3208	−3067	−1441	−922	−714	−13.91	1041	3.6	30600	7.0	3.0	86.04
	6.3	17	−496	−680	−1091	−1888	−3028	−2948	−1532	−990	−742	−16.92	1041	3.6	30600	7.0	3.0	63.37
	7.0	−66	−780	−786	−1172	−1965	−3189	−3044	−1498	−1019	−824	−20.72	1041	3.6	30600	7.0	3.0	8.32
	8.0	−65	−1226	−1369	−1791	−2111	−2990	−2820	−1848	−944	−790	−26.91	1041	3.6	30600	7.0	3.0	0.00
A. saitoi	5.0	640	769	71	−779	−1696	−2146	−1996	−1077	−826	−711	−2.34	973	3.8	12200	6.0	73.0	1.45
proteinase	5.3	866	805	−457	−944	−1463	−2193	−2142	−1115	−819	−794	−3.21	973	3.8	12200	6.0	73.0	0.93
	5.8	154	258	−831	−1471	−2027	−2024	−2063	−1337	−1083	−966	−4.17	973	3.8	12200	6.0	73.0	0.58
	6.3	145	247	−957	−1378	−1821	−2085	−2019	−1244	−972	−821	−5.79	973	3.8	12200	6.0	73.0	0.00
	7.0	15	−722	−1787	−2028	−2319	−2396	−2246	−1465	−1191	−1090	−8.65	973	3.8	12200	6.0	73.0	0.00
	8.0	−503	−1267	−2316	−2737	−2742	−2182	−2082	−1460	−1242	−1148	−13.90	973	3.8	12200	6.0	73.0	0.00
penicillopepsin	5.0	−266	842	519	−231	−796	−2212	−1959	−261	132	252	−32.84	933	3.4	11600	3.0	23.0	22.95
	6.3	−494	214	238	−704	−1352	−2419	−2114	−295	97	211	−39.46	933	3.4	11600	3.0	23.0	0.75
	8.0	−3201	−6340	−6467	−6195	−5208	−3388	−3042	−1794	−1594	−1500	−47.56	933	3.4	11600	3.0	23.0	0.00
trypsin	6.3	−1793	−3050	−3185	−3365	−3560	−2945	−2568	−1412	−1279	−1170	15.77	1034	1.9	9600	12.0	6.0	0.02
chymotrypsin	6.3	4110	870	−6204	−8830	−9730	−7270	−6200	−4260	−4110	−4110	16.27	1030	0.7	9800	5.0	9.0	3.07
papain	6.3	14820	7590	−6990	−8880	−8165	−12580	−11590	−11700	−11700	−11640	19.44	1159	0.6	8700	12.0	19.0	2.17

[a] Wavelengths in nanometers. [b] ZP = ζ potential. [c] H(av) = Bigelow average hydrophobicity. [d] CHG = ratio of acidic to basic amino acid residues. [e] ASA = accessible surface area. [f] ANS = apparent surface hydrophobicity using 1-anilino-8-naphthalenesulfonate. [g] CPA = apparent surface hydrophobicity using *cis*-parinaric acid. [h] MC/PA = milk-clotting to proteolytic activity ratio.

Table II. Variance Explained and Cumulative Proportion of Total Variance Accounted for by Each Factor Derived from the Principal Component Analysis

factor	var explained (eigenvalue)	cum proportn total var	factor	var explained (eigenvalue)	cum proportn total var
1	8.113	0.5071	9	0.061	0.9969
2	2.793	0.6816	10	0.028	0.9987
3	1.867	0.7984	11	0.011	0.9994
4	1.280	0.8783	12	0.006	0.9997
5	0.749	0.9251	13	0.002	0.9998
6	0.562	0.9603	14	0.001	0.9999
7	0.380	0.9840	15	0.001	0.9999
8	0.145	0.9931	16	0.000	1.0000

Table III. Factor Loadings for the Factors Whose Eigenvalues Exceed 1.0

orig variable[a]	factor loading 1	2	3	4
213 nm	0.978	0.000	0.000	0.000
202 nm	0.974	0.000	0.000	0.000
210 nm	0.965	0.000	0.000	0.000
224 nm	0.954	0.000	0.000	0.000
225 nm	0.945	0.000	0.000	0.000
190 nm	−0.905	0.000	0.000	0.000
193 nm	−0.879	0.359	0.000	0.000
H_ϕ(av)	−0.719	0.000	0.000	0.394
ANS	−0.626	0.000	−0.495	−0.250
198 nm	0.000	0.971	0.000	0.000
200 nm	0.000	0.967	0.000	0.000
202 nm	0.537	0.811	0.000	0.000
acidic/basic	0.000	0.000	0.870	0.000
ζ potential	0.000	0.000	−0.692	0.000
A_s	0.000	0.000	0.000	0.794
CPA	0.000	0.000	−0.470	−0.780

[a] Variable abbreviations are similar to those used in Table I.

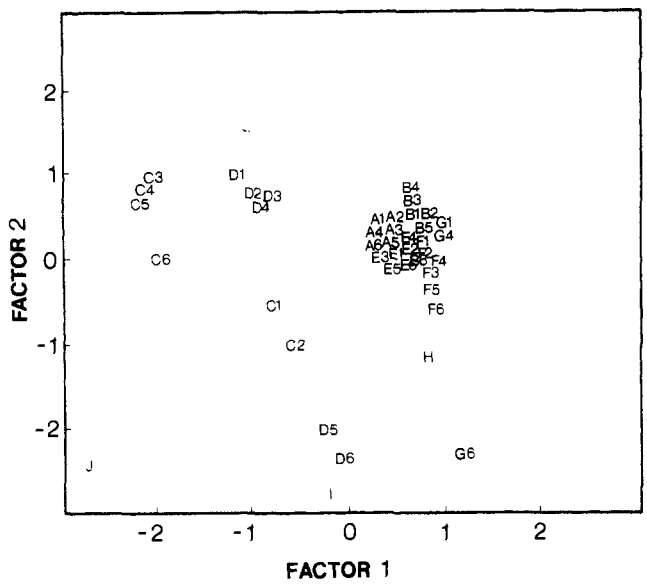

Figure 1. Factor 2 vs. factor 1 obtained from the principal component analysis of the various structural and intrinsic properties of the proteinases: A, *M. miehei* proteinase; B, *E. parasitica* proteinase; C, chymosin; D, pepsin; E, *M. pusillus* proteinase; F, *A. saitoi* proteinase; G, penicillopepsin; H, trypsin; I, α-chymotrypsin; J, papain. The numbers 1–6 represent pH 5.0, 5.3, 5.8, 6.3, 7.0, and 8.0, respectively.

activity ratios. Chymosin at pH 5.0, 5.3, and 8.0, as well as pepsin at pH 7.0 and 8.0, was not associated with the respective groups, which could be explained by noting that these samples showed CD spectral patterns different from those from the previously mentioned pH values (not shown). Since factor 1 and factor 2 were almost exclusively

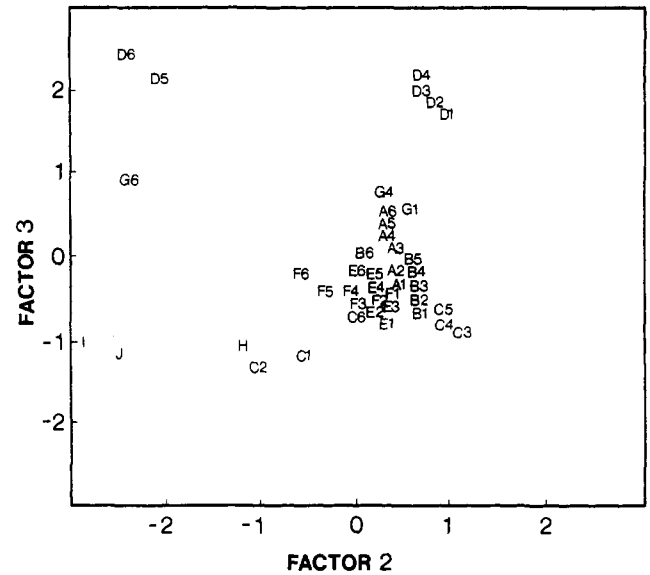

Figure 2. Factor 3 vs. factor 2 obtained from the prinicpal component analysis of the various structural and intrinsic properties of the proteinases: A, *M. miehei* proteinase; B, *E. parasitica* proteinase; C, chymosin; D, pepsin; E, *M. pusillus* proteinase; F, *A. saitoi* proteinase; G, penicillopepsin; H, trypsin; I α-chymotrypsin; J, papain. The numbers 1–6 represent pH 5.0, 5.3, 5.8, 6.3, 7.0, and 8.0, respectively.

concerned with CD spectral data, any change in spectral characteristics of a sample within a group may result in the exclusion of that sample from its group.

The far-UV CD spectra of the aspartyl proteinases examined generally changed at values greater than pH 6.3, which corresponded to the pH at which a substantial decrease in the milk-clotting to proteolytic activity ratio became evident. This change in CD spectra was reflected in a downward shift of the points representing these samples in the plot. The degree of downward shift was related to the degree of change in the CD spectra (i.e., the greater the change in the CD spectra, the larger the downward shift). The shifts or changes observed were in the vertical direction, which would tend to imply that changes occurring in the short-wavelength far-UV CD range were more critical than those seen at the long-wavelength far-UV CD range.

Unlike the plot of factor 2 vs. factor 1, in which both factors were concerned with CD spectral data, the plot of factor 3 vs. factor 2 (Figure 2) was a plot of properties associated with structure (i.e., ζ potential, acidic to basic amino acid ratio, ANS, and CPA hydrophobicity) vs. CD spectral properties (i.e., molar ellipticity values at wavelengths 193, 198, 200, and 202 nm). The plot of factor 3 vs. factor 2 (Figure 2) again showed that the microbial proteinases were closely associated with one another. Chymosin as a group, however, was much more closely associated with the microbial proteinases than in the plot of factor 2 vs. factor 1 (Figure 1). Penicillopepsin at pH 5.0 and 6.3 formed a group intermediate to the microbial enzymes and a group from pepsin at pH 5.0, 5.3, 5.8, and 6.3. Again papain, α-chymotrypsin, and trypsin were not associated with any group.

In Figure 2 it was noted that as the pH of the aspartyl proteinase samples was increased, the points representing those samples tended to shift horizontally to the left-hand side of the plot. Since factor 2 represents CD spectral data and is plotted on the abscissa, any changes in CD spectral data would be expected to cause changes along that axis. The samples corresponding to the points that had been shifted horizontally also had lower milk-clotting to pro-

teolytic activity ratios than those points on the right-hand side of the plot (Table I).

Examination of the milk-clotting to proteolysis activity ratios in relation to the classifications (Figures 1 and 2) generally indicated that proteinases with similar activities were grouped closer together. This would suggest that the structural and intrinsic properties studied partially contribute to enzymatic activity although other factors may also be important. For example, *A. saitoi* proteinase, which had measured characteristics generally similar to those of the microbial proteinases examined, was grouped close to these proteinases but had a very much lower milk-clotting to proteolytic activity ratio.

On the basis of examination of the two-dimensional principal component plots and the analysis of the factor loadings and standardized scores, a manipulation of the original variables may result in a protein with similar structural combination characteristics having similar functional properties. For example, in the plot of factor 3 vs. factor 2 (Figure 2) no change in factor 2 (i.e., CD spectral characteristics) would be necessary to move pepsin at pH values 5.0–6.3 closer to chymosin since the respective groups are situated above one another in the plot. However, a change in factor 3 (representing acidic to basic amino acid ratio, ζ potential, ANS, and CPA hydrophobicity) would be necessary to move pepsin closer to chymosin in Figure 2. Since chymosin had a low acidic to basic amino acid ratio, high ζ potential values and high ANS and CPA hydrophobicity values as compared to pepsin (Table I), hypothetically altering those values for pepsin to more closely resemble those of chymosin would undoubtedly move pepsin closer to chymosin. A similar approach could be used for the other proteinases in the plot. Whether or not such changes could result in enzymatic properties similar to those of a desired enzyme (e.g., chymosin) is speculative and must await further experimental evidence (e.g., chemical modification studies).

Registry No. Aspartic proteinase, 78169-47-8; *M.* aspartic proteinase, 9073-79-4; *E.* aspartic proteinase, 37205-60-0; *A.* aspartic proteinase, 9025-49-4; chymosin, 9001-98-3; pepsin, 9001-75-6; penicillopepsin, 9074-08-2; trypsin, 9002-07-7; chymotrypsin, 9004-07-3; papain, 9001-73-4.

LITERATURE CITED

Aishima, T. *Agric. Biol. Chem.* **1979a**, *43*, 1711–1718.
Aishima, T. *Agric. Biol. Chem.* **1979b**, *43*, 1905–1910.
Aishima, T. *Agric. Biol. Chem.* **1979c**, *43*, 1935–1943.
Arima, K.; Yu, J.; Iwasaki, S. *Methods Enzymol.* **1970**, *XIX*, 446–459.
Arnon, R. *Methods Enzymol.* **1970**, *XIX*, 226–244.
Bech, A.-M.; Foltmann, B. *Neth. Milk Dairy J.* **1981**, *35*, 275–280.
Berridge, N. J. *Biochem. J.* **1945**, *39*, 179–186.
Bigelow, C. C. *J. Theor. Biol.* **1967**, *16*, 187–211.
Chang, C. T.; Wu, C.-S. C.; Yang, J. T. *Anal. Biochem.* **1978**, *91*, 13–31.
Chen, G. C.; Yang, J. T. *Anal. Lett.* **1977**, *10*, 1195–1207.
Cunningham, L. W., Jr.; Tietze, F.; Green, N. M.; Neurath, H. *Discuss. Faraday Soc.* **1953**, *13*, 58–67.
Dalgleish, D. G. *Developments in Dairy Chemistry—1*; Fox, P. F., Ed.; Applied Science: New York, 1982; pp 157–187.
Foltmann, B. *C.R. Trav. Lab. Carlsberg* **1966**, *35*, 143–231.
Foltmann, B. *Neth. Milk Dairy J.* **1981**, *35*, 223–231.
Foltmann, B.; Pedersen, V. B. *Acid Proteases: Structure, Function and Biology*; Tang, J., Ed.; Advances in Experimental Medicine and Biology; Plenum: New York, 1976; Vol. 95, pp 3–22.
Frane, J.; Jennrich, R.; Sampson, P. *BMDP Statistical Software*; Department of Biomathematics, University of California, Los Angeles; University of California Press: Los Angeles, 1981.
Fruton, J. S. "The Enzymes"; Boyer, P. D., Ed.; Academic: New York, 1970; Vol. 3, pp 119–164.
Green, M. L. *J. Dairy Res.* **1972**, *39*, 261–273.
Hayakawa, S.; Nakai, S. *J. Food Sci.* **1985**, *50*, 486–491.
Hsu, I.-N.; Delbaere, L. T. J.; James, M. N. G.; Hofmann, T. *Nature* **1977**, *266*, 140–145.
Ichishima, E. *Methods Enzymol.* **1970**, *XIX*, 397–406.
Ichishima, E.; Yoshida, F. *J. Biochem.* **1966**, *59*, 183–185.
Iwasaki, S.; Tamura, G.; Arima, K. *Agric. Biol. Chem.* **1967**, *31*, 546–551.
Janin, J. *J. Mol. Biol.* **1976**, *105*, 13–14.
Jeffers, J. N. R. *Appl. Stat.* **1967**, *16*, 225–236.
Kwan, K. K. H.; Nakai, S.; Skura, B. J. *J. Food Sci.* **1983**, *48*, 1418–1421.
Laemmli, U. K. *Nature* **1970**, *227*, 680–685.
Ottesen, M.; Rickert, W. *Methods Enzymol.* **1970**, *XIX*, 459–460.
Schulz, G. E.; Barry, C. D.; Friedman, J.; Chou, P. Y.; Fasman, G. D.; Finkelstein, A. V.; Lim, V. I.; Ptitsyn, O. B.; Kabat, E. A.; Wu, T. T.; Levitt, M.; Robson, B.; Nagano, K. *Nature* **1974**, *250*, 140–142.
Smith, E. L.; Kimmel, J. R. *The Enzymes*; Boyer, P. D., Lardy, H., Myrback, K., Eds.; Academic: New York, 1960; Vol. 4, pp 133–173.
Sober, H. A. *Handbook of Biochemistry*, 2nd ed.; Chemical Rubber Co.: Cleveland, 1970.
Sodek, J.; Hofmann, T. *Methods Enzymol.* **1970**, *XIX*, 391–397.
Stuper, A. J.; Brugger, W. E.; Jurs, P. C. *Computer Assisted Studies of Chemical Sturcture and Biological Functions*; Wiley: New York, 1979.
Townsend, A.-A.; Nakai, S. *J. Food Sci.* **1984**, *48*, 588–594.
Visser, S. *Neth. Milk Dairy J.* **1981**, *35*, 65–88.
Voynick, I. M.; Fruton, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 257–259.
Walsh, K. A.; Neurath, H. *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *52*, 884–889.
Whitaker, J. R. *Methods Enzymol.* **1970**, *XIX*, 436–445.

Received for review September 30, 1985. Revised manuscript received February 27, 1986. Accepted April 3, 1986.