BIOCHEMICAL JOURNAL LETTERS

On the structure/function relationship of polymorphonuclear-leucocyte collagenase

In their recent paper, Knauper et al. [1] discuss the autoproteolytic degradation of polymorphonuclear-leucocyte (PMNL) collagenase. They were able to show that the enzyme is processed to two major fragments of 40 kDa and 27 kDa, corresponding to the enzymic core and to the hemopexin-like C-terminal domain respectively. By sequencing the N-terminus of each fragment, it was possible to deduce the autoproteolytic cleavage point, the peptide bond between either Gly^{242}-Leu^{243} or Pro^{247}-Ile^{248}. As expected, both fragments lack the ability to cleave interstitial collagen. More interesting, however, was the fact that both fragments also lacked the ability of binding to collagen.

Our group has suggested that the sequence ^{244}SSNPIQP^{250} in PMNL collagenase (or SQNPVQP in fibroblast collagenase) is important for its substrate-binding activity [2]. Using chimeric constructions, Hirose et al. [3] have confirmed that this region is important for the collagenolytic activity of the PMNL enzyme. In the Discussion of their paper, Knauper et al. [1] concluded that the sequence SSNPIQP would not be involved in collagen-binding activity, since after autoproteolytic cleavage the ensuing 40 kDa and 27 kDa fragments do not bind collagen. However, they did not notice that the cleavage point is exactly in the middle of the SSNPIQP sequence. In our opinion, the fact that the resulting fragments do not bind to collagen, whereas the intact molecule does, supports our suggestion that the sequence SSNPIQP is involved in collagen binding, rather than the opposite, as they contend. It seems clear, however, that the collagen-binding capacity of collagenases involves more than one peptide stretch. A great deal of evidence seems to indicate that additional sequences present in the hemopexin-like C-terminal domain are also involved in collagen-binding activity. For instance, mutants of neutrophil collagenase lacking a part of this domain show a very weak collagenolytic activity [3]. One interesting possibility is that these sequences are responsible for an 'unproductive binding' and that the SQNPVQP- or SSNPIQP-mediated binding could adjust the enzyme in the correct place. In agreement with this possibility is the fact that fibroblast collagenase chimaeras containing the C-terminal domain of stromelysin (and without the SQNPVQP sequence) also bind collagen and likewise lack the ability to cleave it [4]. The same is true for the native stromelysin.

Knauper et al. [1] remind us that autoproteolysis in collagenases takes place in a proline-rich region and that the cleavage point (Gly-Leu or Pro-Ile) resembles that of collagen. It is well known from Ramachandran's work that proline-rich regions can adopt a collagen-like conformation [5]. It is attractive to speculate that the proline-rich region of collagenases has a collagen-like conformation. Because the collagenase-sensitive site in collagen has a more relaxed triple helix, the collagen helix conformation of the proline repeat could enable the enzyme to bind to collagen. This would explain the fact that this region is always the target of autoproteolytic cleavage. If confirmed, this would be one of the first cases where an enzyme mimics its substrate's conformation in order to cleave it.

Sandro J. DE SOUZA and Ricardo R. BRENTANI

Ludwig Institute for Cancer Research, R. Antonio Prudente 104, 4° andar, 01509-010 São Paulo, Brasil

1 Knauper, V., Osthues, A., DeClerck, Y. A., Langley, K. E., Blaser, J. and Tschesche, H. (1993) Biochem. J. **291**, 847–854
2 Souza, S. J. and Brentani, R. R. (1992) J. Biol. Chem. **267**, 13763–13767
3 Hirose, T., Patterson, C., Pourmottabed, T., Mainardi, C. L. and Hasty, K. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 2569–2573
4 Murphy, G., Allan, J. A., Willenbrock, F., Cocket, M., O'Connel, J. and Docherty, A. J. P. (1992) J. Biol. Chem. **267**, 9612–9618
5 Ramachandran, G. N. (1967) Chemistry of Collagen, Academic Press, New York

Received 13 September 1993

Protochlorophyllide reductase is homologous to human carbonyl reductase and pig 20β-hydroxysteroid dehydrogenase

Protochlorophyllide reductase is an NADPH-dependent enzyme that catalyses the conversion of protochlorophyllide into chlorophyllide [1]. This is the first detectable light-dependent step towards the formation of the chloroplast. Synthesis and degradation of this enzyme are tightly regulated, due to its important role in the synthesis of chlorophyll in plants. The sequences of the enzymes from barley, oat, pea, pine and *Arabidopsis* have been determined without any reported homology to other enzymes [2–6]. My interest in the connection between animal and plant dehydrogenases [7,8] led me to compare the sequence of protochlorophyllide reductase with the protein database, which revealed that this enzyme is homologous to pig 20β-hydroxysteroid dehydrogenase [9] and human carbonyl reductase [10], enzymes that belong to a protein superfamily that contains human 11β-hydroxysteroid dehydrogenase, 17β-hydroxysteroid dehydrogenase, 15-hydroxyprostaglandin dehydrogenase and *Drosophila melanogaster* alcohol dehydrogenase, as well as bacterial enzymes that are important in synthesis of antibiotics [11–15]. Much is known about these enzymes concerning residues that are important in binding of the nicotinamide cofactor [16], in preference for NADH and NADPH [17,18], and in catalysis [19–22], information that should be useful in understanding catalysis by protochlorophyllide reductase.

As noted by Spano et al. [6] the amino acid sequences of the different protochlorophyllide reductases are very similar to each other: the loblolly-pine enzyme is 76 % identical with either barley or oat, 84 % identical with pea, and 82 % identical with *Arabidopsis* protochlorophyllide reductase. Thus, a comparison of one protochlorophyllide reductase sequence with that of other oxidoreductases is representative of that for other protochlorophyllide reductases. It is an ~ 140-residue segment in exon 3 of protochlorophyllide reductase that is most similar to human

Table 1 Comparison of *Arabidopsis* protochlorophyllide reductase with various oxidoreductases

Similarity between proteins was analysed with the ALIGN program [23]. For the analyses reported here, 1000 random permutations were used for the statistical analysis, and the Dayhoff matrix was used wtih a bias of 6 and a gap penalty of 8. The alignment score is the number of S.D. by which the maximum score for the real sequences exceeds the average maximum score for the random. The probability of getting a score of 9 S.D. by chance is 10^{-19}. Although the score between protochlorophyllide reductase and *Drosophila* alcohol dehydrogenase of 4.9 S.D. ($P = 10^{-6}$) is not sufficient to establish homology, other analyses show that alcohol dehydrogenase is homologous to human carbonyl reductase and human 17β-hydroxysteroid dehydrogenase.

Protein	ALIGN score (S.D. units)
Human carbonyl reductase (residues 9–141)	10.15
Pig 20β-hydroxysteroid dehydrogenase (residues 9–140)	9.5
Human 17β-hydroxysteroid dehydrogenase (residues 5–145)	8.65
Bradyrhizobium japonicum FixR (residues 40–178)	8.5
Streptomyces hydrogenans 20β-hydroxysteroid dehydrogenase (residues 9–130)	7.85
Streptomyces coelicolor β-ketoreductase (residues 10–146)	6.3
Drosophila alcohol dehydrogenase (residues 9–123)	4.9

carbonyl reductase and pig 20β-hydroxysteroid dehydrogenase and their homologues. As seen in Table 1, an ALIGN analysis [23] comparing *Arabidopsis* protochlorophyllide reductase with human carbonyl reductase and pig 20β-hydroxysteroid dehydrogenase yields scores that are more than 9 S.D. above that found in comparisons of random sequences of these proteins. The probability of getting a score of 9 S.D. by chance is 10^{-19}, which indicates that these proteins are homologues, that is, they are descended from a common ancestor.

Protochlorophyllide reductase contains three glycine residues near the N-terminus, with a spacing of Gly-Xaa-Xaa-Xaa-Gly-Xaa-Gly, like that in its homologues [11–13,24]. Mutagenesis of the first glycine in this part of *Drosophila* alcohol dehydrogenase indicates that this residue is critical for binding NAD^+ [16]. Moreover, the tertiary structure of rat dihydropteridine reductase co-crystallized with NADH [24] shows that the glycine motif is part of the adenosine-binding domain. Taken together, this provides experimental support for previous hypotheses [3,6] that the glycine motif is part of the nucleotide-binding domain on protochlorophyllide reductase.

Many oxidoreductases contain a $\beta\alpha\beta$ fold that binds the AMP part of the nucleotide cofactor [25,26]. The $\beta\alpha\beta$ fold appears to be near the amino terminus in homologues of protochlorophyllide reductase [24,27]. It is the residues at the C-terminal part of the second β strand in the $\beta\alpha\beta$ fold that are important in determining preference for NADH or NADPH [25,26,28]. For example, coulombic repulsion between a negatively charged residue, such as aspartic acid, and the 2′-phosphate on NADPH leads to a preference for NADH [25,26,28]. Mutagenesis of aspartic acid-38 in *Drosophila* alcohol dehydrogenase [17] and aspartic acid-37 in human dihydropteridine reductase [18] to replace these residues with an uncharged residue confirms that this part is important in determining specificity for NADPH and NADH.

All three enzymes in Figure 1 are NADPH-dependent. Alanine-36 in carbonyl reductase and alanine-37 in pig 20β-hydroxysteroid dehydrogenase align with the aspartic acid at the end of the second β strand in NADH-dependent enzymes [24]. Cysteine-120 of *Arabidopsis* protochlorophyllide reductase aligns with alanine-36 of carbonyl reductase (Figure 1). The other protochlorophyllide reductases have a cysteine residue at this position, as well as an adjacent arginine that would be expected to attract the 2′-phosphate of NADPH [24,28]. Interestingly, chemical modification of cysteine residues on protochlorophyllide reductase inactivates the enzyme [32].

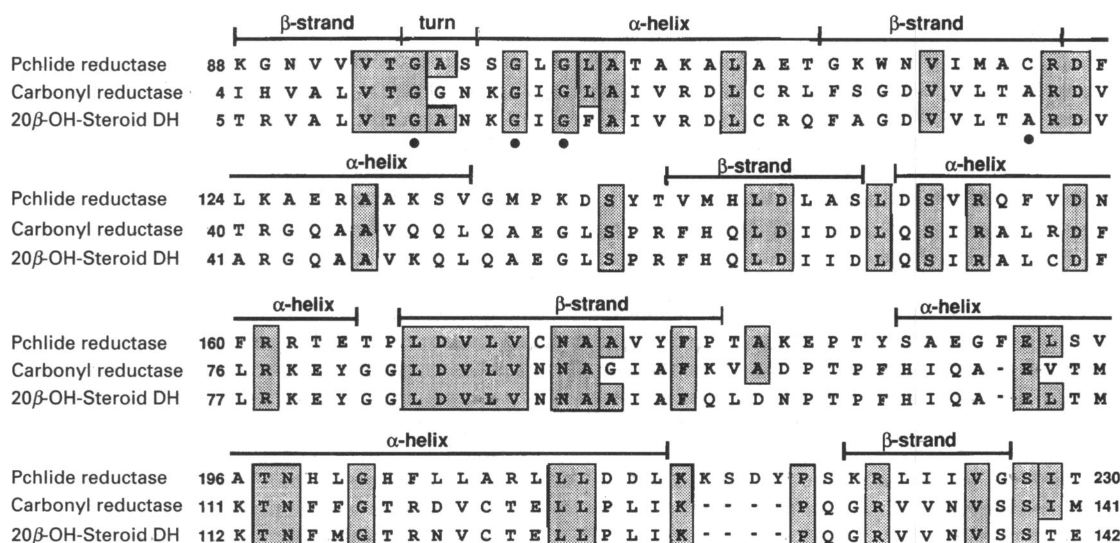

Figure 1 Alignment of *Arabidopsis* protochlorophyllide reductase with human carbonyl reductase and pig 20β-hydroxysteroid dehydrogenase

The algorithm of Feng and Doolittle [29] was used to construct the alignment. In the sequence shown, carbonyl reductase and 20β-hydroxysteroid dehydrogenase (20β-OH-Steroid DH) are 85% identical with each other and about 30% identical with *Arabidopsis* protochlorophyllide reductase (Pchlide reductase). Identities between protochlorophyllide reductase and the other two enzymes are shown in shaded boxes. The secondary structure determined from the tertiary structure of *Streptomyces hydrogenans* 20β-hydroxysteroid dehydrogenase [30,31] and rat dihydropteridine reductase [27] is shown above the alignment. Residues 88–121 of protochlorophyllide reductase are proposed to have a $\beta\alpha\beta$ structure found in the nucleotide-binding domain of oxidoreductases. The symbol ● identifies three glycine residues and cysteine-120 in protochlorophyllide reductase that are proposed to be important in binding the nucleotide cofactor.

Previous sequence analyses [11–13] found that residues corresponding to serine-142, tyrosine-155 and lysine-159 of 17β-hydroxysteroid dehydrogenase are highly conserved in other members of the oxidoreductase superfamily. Mutagenesis studies with human 15-hydroxyprostaglandin dehydrogenase [21], *Drosophila* alcohol dehydrogenase [19,20] and 11β-hydroxysteroid dehydrogenase [22] indicate that this tyrosine and lysine are at the catalytic site. The *Arabidopsis* protochlorophyllide reductase sequence contains a sequence Lys-Ala-Tyr-Lys-Asp-Ser-Lys beginning at residue 274. We suggest that tyrosine-276 and lysine-280 are part of the catalytic site. In homologues of protochlorophyllide reductase, this lysine is thought to lower the pK_a of the nearby tyrosine [20,30] and to orient the nicotinamide ring for pro-*S* hydrogen transfer [32a].

Although the pentapeptide is highly conserved among the protochlorophyllide reductases, there is little similarity to that in homologous enzymes [11–15]. An important difference is that protochlorophyllide reductase contains lysine-274, lysine-277 and aspartic acid-278 close to the proposed catalytically active tyrosine. Thus there are three lysine residues and an aspartic acid that could interact with the cofactor, tyrosine-276, and the substrate, suggesting some differential chemical interactions between protochlorophyllide reductase and its cofactor and substrate, compared with its homologues.

This proposed location of the catalytically active tyrosine means that there is a 47-residue spacing between the conserved serine-228 and tyrosine-276. The extra residues are encoded in exon 4 in pea and pine protochlorophyllide reductase. Almost all homologues have a 12-residue spacing between these serine and tyrosine residues [11–13]. Interestingly, the two exceptions are human carbonyl reductase and pig 20β-hydroxysteroid dehydrogenase, the enzymes that are most similar to protochlorophyllide reductase (Table 1). In human carbonyl reductase and pig 20β-hydroxysteroid dehydrogenase, there are 53 residues between the conserved serine and tyrosine residues.

Homologous oxidoreductases in *Arabidopsis*, *Brassica napus* [33] and corn [34] do not contain the extra 35 residues between their conserved serine and tyrosine residues. Thus plants and animals contain oxidoreductases with both spacings between the conserved serine and tyrosine residues. The most parsimonious interpretation is that a subfamily of oxidoreductases with an additional segment in the exon that contains the catalytic tyrosine arose before the divergence of plants and animals over 1 billion years ago. It will be interesting to see if homologues in unicellular organisms have this extra segment.

The ancestral bacterial homologues of protochlorophyllide reductase include ribitol dehydrogenase, glucitol-6-phosphate dehydrogenase, and β-ketoreductases important in synthesis of polyketide antibiotics. A distinguishing functional feature of these enzymes is they act on secondary alcohols as substrates. For example, *Drosophila* alcohol dehydrogenase has a $k_{cat.}/K_m$ for secondary alcohols such as propan-2-ol and butan-2-ol that is about 10-fold higher than that for propan-1-ol and butan-1-ol [20]. We propose the term *sec*-alcohol dehydrogenase to describe these enzymes, because this name provides functional information.

Received 21 October 1993

I thank Dr. K. I. Varughese, Dr. J. M. Whiteley and Dr. N. H. Xuong for discussion of their research on the catalytic site of dihydropteridine reductase.

Michael E. BAKER

Department of Medicine, 0623B, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0623, U.S.A.

1 Griffiths, W. T. (1974) FEBS Lett. **49**, 196–200

2 Schulz, R., Steinmujller, K., Klaas, M., Forreiter, C., Rasmussen, S., Hiller, C. and Apel, K. (1989) Mol. Gen. Genet. **217**, 355–361

3 Darrah, P. M., Kay, S. A., Teakle, G. R. and Griffiths, W. T. (1990) Biochem. J. **265**, 789–798

4 Benli, M., Schulz, R. and Apel, K. (1991) Plant Mol. Biol. **16**, 615–625

5 Spano, A. J., He, Z., Michel, H., Hunt, D. F. and Timko, M. P. (1992) Plant Mol. Biol. **18**, 967–972

6 Spano, A. J., He, Z. and Timko, M. P. (1992) Mol. Gen. Genet. **236**, 86–95

7 Baker, M. E. (1992) Can. J. Microbiol. **38**, 541–547

8 Baker, M. E. (1992) J. Steroid Biochem. Mol. Biol. **41**, 301–308

9 Tanaka, M., Ohno, S., Adachi, S., Nakajin, S., Shinoda, M. and Nagaham, Y. (1992) J. Biol. Chem. **267**, 13451–13455

10 Wermuth, B. (1992) Prostaglandins **44**, 5–9

11 Tannin, G. M., Agarwal, A. K., Monder, C., New, M. I. and White, P. C. (1991) J. Biol. Chem. **266**, 16653–16658

12 Persson, B., Krook, M. and Jornvall, H. (1991) Eur. J. Biochem. **200**, 537–543

13 Krozowski, Z. (1992) Mol. Cell. Endocrinol. **84**, C25–C31

14 Baker, M. E. (1991) Prostaglandins **42**, 391–407

15 Baker, M. E. (1991) Steroids **56**, 354–360

16 Chen, Z., Lu, L., Shirley, M., Lee, W. R. and Chang, S. H. (1990) Biochemistry **29**, 1112–1118

17 Chen, Z., Lin, Z.-G., Lee, W. R. and Chang, S. H. (1991) Eur. J. Biochem. **202**, 263–267

18 Grimshaw, C. E., Matthews, D. A., Varughese, K. I., Skinner, M., Xuong, N. H., Bray, T., Hoch, J. and Whiteley, J. M. (1992) J. Biol. Chem. **267**, 15334–15339

19 Albalat, R., Gonzalez-Duarte, R. and Atrian, S. (1992) FEBS Lett. **308**, 235–239

20 Chen, Z., Jiang, J. C., Lin, Z. G., Lee, W. R., Baker, M. E. and Chang, S. H. (1993) Biochemistry **32**, 3342–3346

21 Ensor, C. M. and Tai, H.-H. (1991) Biochem. Biophys. Res. Commun. **176**, 840–845

22 Obeid, J. and White, P. C. (1992) Biochem. Biophys. Res. Commun. **188**, 222–227

23 Dayhoff, M. O., Barker, W. C. and Hunt, L. T. (1983) Methods Enzymol. **91**, 524–545

24 Baker, M. E. (1994) Steroids **59**, 248–258

25 Wierenga, R. K., De Maeyer, M. C. and Hol, W. G. J. (1985) Biochemistry **24**, 1346–1357

26 Branden, C. and Tooze, J. (1991) Introduction to Protein Structure, Garland Publishing, New York

27 Varughese, K. I., Skinner, M. M., Whiteley, J. M., Matthews, D. A. and Xuong, N. H. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 6080–6084

28 Scrutton, N. S., Berry, A. and Perham, R. N. (1990) Nature (London) **343**, 38–43

29 Feng, D. and Doolittle, R. F. (1990) Methods Enzymol. **183**, 375–387

30 Ghosh, D., Weeks, C. M., Groschulski, P., Duax, W. L., Erman, M., Rimsay, R. L. and Orr, J. C. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 10064–10068

31 Krook, M., Ghosh, D., Stromberg, R., Carlquist, M., and Jornvall, H. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 502–506

32 Oliver, R. P. and Griffiths, W. T. (1981) Biochem. J. **195**, 93–101

32a Varughese, K. I., Xuong, N. H., Kiefer, P. M., Matthews, D. A. and Whiteley, J. M. (1994) Proc. Natl. Acad. Sci. U.S.A., in the press

33 Slabas, A. R., Chase, D., Nishida, I., Murata, N., Sidebottom, C., Safford, R., Sheldon, P. S., Kekwick, R. G. O., Hardie, D. G. and Mackintosh, R. W. (1992) Biochem. J. **283**, 321–326

34 DeLong, A., Calceron-Urrea, A. and Dellaporta, S. L. (1993) Cell **74**, 757–768