FEBS Letters 407 (1997) 32–36 FEBS 18447

Scavengase p20: a novel family of bacterial antioxidant enzymes

Xiao-Yu Wan[1], Yuan Zhou[2], Zi-Ying Yan, Hai-Lin Wang, Yun-De Hou, Dong-Yan Jin*

National Key Laboratory of Molecular Biology and Genetic Engineering, Chinese Academy of Preventive Medicine, 100 Yingxin Street, Beijing 100052, People's Republic of China

Received 30 January 1997; revised version received 4 March 1997

Abstract A novel antioxidant enzyme designated scavengase p20 was identified in various pathogenic bacteria through database searching for sequences strikingly homologous to a recently discovered *Escherichia coli* thiol peroxidase p20. The direct biochemical evidence for the existence of scavengase p20 in *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Helicobacter pylori* was provided by protein microsequencing and by in vitro assays for antioxidant activities. Overlapping genes encoding scavengase p20 and superoxide dismutase were recognized in *H. pylori* and their functional implications are discussed.

Key words: Scavengase p20; Antioxidant enzyme; Thioredoxin-dependent peroxidase; *Helicobacter pylori*; Sequence analysis; Protein microsequencing

1. Introduction

The oxidative burst is one of the major mechanisms by which leukocytes kill pathogenic microorganisms. To survive this killing effect, pathogens have developed many antioxidant enzyme systems to detoxify reactive oxygen species (ROS). The imbalance of host redox conditions as a result of microbial infection is believed to play an important role in the pathogenesis of various diseases [1–4].

Well-documented bacterial antioxidant enzymes include superoxide dismutase (SOD), catalase and alkyl hydroperoxide reductase (AhpR). SOD catalyses the dismutation of superoxide anion to hydrogen peroxide [4–6], which is in turn broken down by catalase and other peroxidases. AhpR is a glutathione-independent peroxidase system with two subunits, AhpF and AhpC [7,8]. AhpC was recently shown to belong to the thioredoxin-dependent peroxidase family called peroxiredoxin [7,9].

Besides all these known bacterial antioxidants, a novel thioredoxin-linked thiol peroxidase p20 was recently identified from the periplasmic space of *Escherichia coli* [10]. Interestingly, this protein displays antioxidant activities to prevent glutamine synthetase (GS) from inactivation by the metal-catalysed thiol system and to scavenge H_2O_2 in the presence of the thioredoxin system, but it shares no significant homology with any known peroxidase [10]. This enzyme was originally named 'thiol peroxidase'. This nomenclature is confusing since some of the peroxidases in the peroxiredoxin family have already been designated 'thiol-specific antioxidant (TSA)' [7,11] and they do possess thiol peroxidase activity [12]. *E. coli* p20 and TSAs may act through similar mechanisms, but they are structurally distinct. With this in mind, we propose to rename p20 as 'scavengase'. The antioxidant activity of scavengase p20 to scavenge H_2O_2 is well-characterized [10]. However, the presence of scavengase p20 in other microorganisms is not described and the functional implications of scavengase p20 in bacterial infection and pathogenesis are poorly understood.

To obtain some insight into the distribution and biological functions of scavengase p20, here we performed a database search for protein sequences strikingly homologous to *E. coli* p20 and then provided biochemical characterization of scavengase p20 from *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Helicobacter pylori*. Additionally, we described a previously unidentified overlapping gene in *H. pylori* that codes for scavengase p20 and SOD.

2. Materials and methods

2.1. Sequence analysis

Nucleotide and peptide sequences were analyzed with the help of the Wisconsin software package (Version 8.1, Genetic Computer Group, Inc., Madison, WI). Database searching was performed using the BLAST programs [13]. Multiple alignment of protein sequences was generated with a progressive pairwise algorithm [14]. Phylogenetic analysis was based on a matrix of evolutionary distances and the phylogeny was reconstructed with the UPGMA method [15].

2.2. Bacterial strains

E. coli K-12 RR1, *H. influenzae* Rd, *H. pylori* NCTC 11639 and *S. pneumoniae* R36A were originally obtained from American Type Culture Collection (ATCC, Rockville, MD) and grown as recommended by ATCC.

2.3. Protein purification

Bacterial cell extracts were prepared by sonication/centrifugation and scavengase was partially purified through 70% ammonium sulfate precipitation, DEAE-cellulose (×1) and Sephadex G75 (×2) column chromatography. The purification was carried out according to a published protocol with minor modifications [10]. The target protein was traced by its GS protection activity. The yield, purity and purification factor (fold enhancement of the product) were calculated for each preparation and were shown to be reproducible for each bacterium. The purification product of >70% purity as examined by SDS-PAGE (see below, Fig. 3) was used as 'purified scavengase p20' throughout this work.

2.4. Protein microsequencing

Protein samples run on the unstained SDS-PAGE gel (using Tris-Tricine as electrophoresis buffer) were electroblotted onto an Immobilon-P polyvinylidene difluoride (PVDF) membrane (using 10 mM 3-cyclohexylamino-1-propanesulfonic acid (CAPS) with 10% methanol as transfer buffer). The PVDF membrane was stained with 0.1% Coomassie blue R250 for 30 s and washed with 50% methanol. Pieces of membrane with the 20 kDa protein bands were trimmed to 3×7 mm

*Corresponding author. Fax: (86) (10) 63541221.

[1]*Present address:* Johns Hopkins University Oncology Center, Baltimore, MD 21231, USA.

[2]*Present address:* Department of Biology, Georgetown University, Washington, DC 20057, USA.

PII S0014-5793(97)00302-5

size and subjected to microsequencing with Edman degradation as previously described [16]. Sequencing was automatically performed on an Applied Biosystems protein sequencer (model 477A) coupled to an on-line amino acid analyzer (model 120A).

2.5. Antioxidant assays

The GS protection assay was carried out at 37°C essentially as described [9,10,17]. The reaction mixture (100 μl) contained 5 μg *E. coli* GS (Sigma, St. Louis, MO), 10 μM $FeCl_3$, 10 mM dithiothreitol (DTT), 10 μg purified scavengase p20 (prepared as described above) and 100 mM HEPES (pH 7.0). Aliquots (15 μl) of reaction product were removed at certain time points and assayed for GS activity.

Indirect peroxidase assay was performed as described elsewhere [10,17]. The reaction mixture (0.5 ml) contained *E. coli* thioredoxin reductase (20 μg), *E. coli* thioredoxin (20 μg), H_2O_2 (5 mM), purified scavengase p20 (25 μg) and NADPH (0.25 mM). NADPH oxidation coupled to H_2O_2 reduction was monitored as absorbance units at 340 nm (AU_{340}). Peroxidase activity was expressed as AU_{340}/min.

3. Results and discussion

3.1. Identification of the family of scavengase p20

We searched the sequence database for scavengase p20 homologs with the help of the BLAST programs. Surprisingly, we found that scavengase p20 is widely distributed in various pathogenic bacteria including *H. influenzae* [18], *Vibrio cholerae* [19], *S. pneumoniae* [20], *S. sanguis* [21,22], *S. gordonii* [23], *S. parasanguis* [21] and *H. pylori* [6]. Meanwhile, no eukaryotic homolog was identified. A sequence alignment (Fig. 1) and a phylogenetic tree (Fig. 2) were generated for scavengase p20 from different species. Both clearly indicate that these proteins are the founding members of a novel family of antioxidants. The level of homology between scavengase p20 proteins from different bacteria is high (37–63% identity and 57–73% similarity) and their predicted molecular size is similar (about 20 kDa). From the consensus sequence summarized below the alignment (Fig. 1), conserved residues including a cysteine can be found.

It is noteworthy that scavengase p20 from many pathogens was originally described as an important determinant for virulence. *V. cholerae* scavengase p20 was first identified as TagD for ToxR-activated gene D. TagD is a member of the toxin-coregulated pilus (TCP)-encoding gene cluster (*tcp*). It is known that the type-IV pilus specified by the *tcp* cluster is a major colonization determinant [19]. *Streptococcus* sp. scavengase p20 genes have also been hypothesized to be members of an operon/gene cluster encoding adhesins [20–23]. However, further studies are required to elucidate the biological functions of scavengase p20 and the roles it plays in bacterium-host interaction.

3.2. Biochemical characterization of scavengase p20

To further characterize scavengase p20 from *E. coli*, *H.*

```
                1                                                   50
S.gordonii      ...MTTFLGN PVTFTGKQLQ VGDTAHDFSL TATDLSKKTL ADFAGKKKVL
S.sanguis       ...MTTFLGN PVTFTGKQLQ VGDTAHDFSL TATDLSKKTL ADFAGKKKVL
S.pneumoniae    ...MTTFLGN PVTFTGSQLQ VGDTAHDFSL TTPNLEKKSL ADFAGKKKVL
S.parasanguis   ....ATFLGN PVTFTGSQLQ VGEIAHDFSL ITPALEKKSL ADFAGKKKVL
E.coli          MSQTVHFQGN PVTVANSIPQ AGSKAQTFTL VAKDLSDVTL GQFAGKRKVL
H.influenzae    ..MTVTLAGN PIEVGGHFPQ VGEIVENFIL VGNDLADVAL NDFASKRKVL
V.cholerae      ..MTVTFQNN PVSISGSFPK VGDRLPSFTL CGADLNDLNN EDFKGKKIVM
H.pylori        ......FKEE TYQLEGKPLK VGDKAPDVKL VNGDLQEVNL LKQGVRFQVV
                                     *       *     *              *

                51                                                  100
S.gordonii      SIIPSIDTGV CSTQTRRFNQ ELSDLDNTVV ITVSVDLPFA QGKWCAAEGI
S.sanguis       SIIPSIDTGV CSTQTRRFNQ ELSDLDNTVV ITVSVDLPFA QGKWCAAEGI
S.pneumoniae    SVI
S.parasanguis   SIIPSIDTGI CSMQTRHFNK TLSDLEDTVV LTVSVDLPFA QGKWCAAEGL
E.coli          NIFPSIDTGV YAASVRKFNQ LATEIDNTVV LCISADLPFA QSRFCGAEGL
H.influenzae    NIFPSIDTGV CATSVRKFNQ QAAKLSNTIV LCISADLPFA QARFCGAEGI
V.cholerae      SIFPSIDTPV CSKSVKVLQN ALMTRSDTVL LCVSADLPFA MSRFCTEHAV
H.pylori        SALPSLTGSV CLLQAKHFNE QAGKLPSVSF SVISMDLPFS QGQICGAEGI
                  **                                * ****      *

                101                                                 150
S.gordonii      ENAVMLSDYF DHSFGRDYAV LINE...WHL LARAVLVLDE NNTVTYAEYV
S.sanguis       ENAVMLSDYF DHSFGRDYAV LINE...WHL LARAVLVLDE NNTVTYAEYV
S.parasanguis   DNAIMLSDYY DHSFGKAYGL LINE...WHL LARAVLVLDA DNKITYVEYL
E.coli          NNVITLSTFR NAEFLQAYGV AIADGPLKGL AARAVVVIDE NDNVIFSQLV
H.influenzae    ENAKTVSTFR NHALHSQLGV DIQTGPLAGL TSRAVIVLDE QNNVLHSQLV
V.cholerae      ANVTNASFFR EPAFTERFGV NLNEGALRGL AARAVIVADE FGVITHSELV
H.pylori        KDLRILSDFR YKAFGENYGV LLGKGSLQGL LARSVFVLDD KGVVIYKEIV
                      *                           *   * * * *

                151             170
S.gordonii      DNINTEPDYD AAIAAVKNL
S.sanguis       DNINTEPDYD AAIAAVKSL
S.parasanguis   DNINSEPNYD AAIEAVKVLG
E.coli          DEITTEPDYE AALAVLKA
H.influenzae    EEIKEEPNYE AALAVLA
V.cholerae      NEITNEPDYD RILMSL
H.pylori        QNILEEPNYE VLLKVLE
                  *  **
```

Fig. 1. Amino acid sequence alignment of the p20 scavengase proteins from different species. The consensus sequence is highlighted by asterisks (*). The peptide segments which have been sequenced by the Edman degradation method are underlined. Listed below are GenBank accession numbers of the sequences being compared: *S. gordonii*, L11577; *S. sanguis*, M63481; *S. pneumoniae* (partial sequence), L19055; *S. parasanguis*, M26130; *E. coli*, U33213; *H. influenzae*, U32705; *V. cholerae*, U02375; *H. pylori*, X72618.

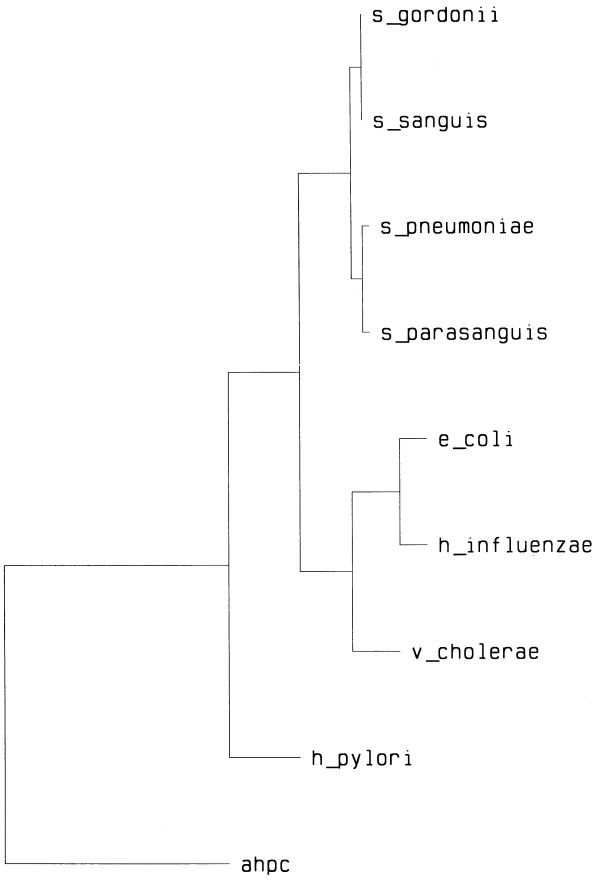

Fig. 2. Distance matrix tree for the scavengase p20 family. An *E. coli* AhpC protein (ahpc) is also shown as a reference.

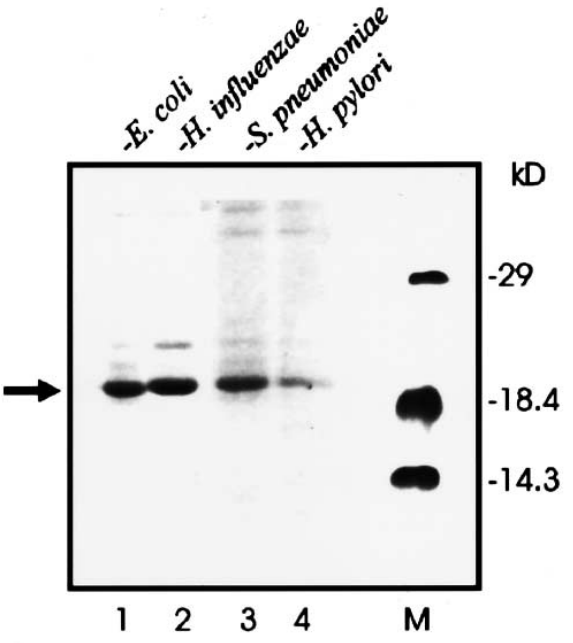

Fig. 3. SDS-PAGE (15%) analysis of scavengase p20 preparations. Shown on the right are the positions of molecular weight markers. The band corresponding to scavengase p20 is highlighted by an arrow.

influenzae, *H. pylori* and *S. pneumoniae*, we prepared purified enzymes from these bacteria and analyzed the preparations by SDS-PAGE. As shown in the electrophoretogram (Fig. 3) of the Coomassie blue R250-stained gel, a 20 kDa protein of $>70\%$ purity is evident when 10 μg of purified scavengase p20 was loaded onto each lane. The same samples run in parallel on the unstained half of the gel were electroblotted onto PVDF membrane and then sequenced directly by Edman degradation. The N-terminal amino acid sequence of the 20 kDa protein from *H. influenzae* and *S. pneumoniae* thus determined is underlined in Fig. 1. This sequence was identical to that predicted from DNA [18,20]. The first few amino acid residues from the 20 kDa protein in the *H. pylori* extract were recognized as (X)-(X)-Val-Thr-Phe-Lys-Glu-Glu upon careful analysis of the microbore-HPLC chromatograms of the phenylthiohydantoin (PTH)-amino acids. This is suggestive of a sequencing error in the published *H. pylori* DNA sequence (EMBL nucleotide sequence database number X72618) [6]. Regardless of this discrepancy, our microsequencing data provide direct evidence that scavengase p20 does exist in *H. influenzae*, *S. pneumoniae* and *H. pylori*.

It is clear from the SDS-PAGE and the microsequencing results that the major component of our preparations is scavengase p20. The 22 kDa (or 24 kDa) AhpC was undetected. We also note that our preparations of purified scavengase p20 did not react with a rabbit polyclonal antibody raised against AhpC (data not shown). Hence, we went on to test the antioxidant activities of these preparations. As mentioned earlier, scavengase p20 was followed by its GS protection activity during the process of purification. Therefore it is no wonder that the preparations of purified scavengase p20 from *H.*

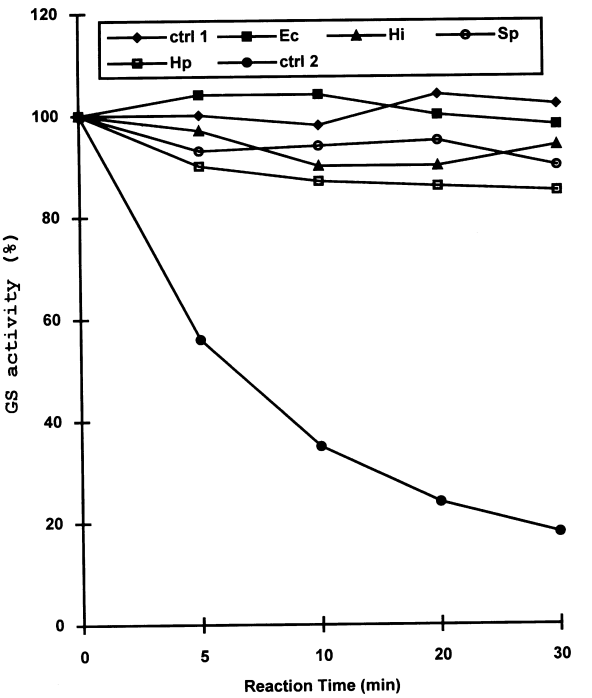

Fig. 4. GS protection activity of scavengase p20 from *E. coli* (Ec), *H. influenzae* (Hi), *S. pneumoniae* (Sp) and *H. pylori* (Hp). 1 mM EDTA was added to control 1 (ctrl 1) to chelate the catalyst Fe^{3+}. Control 2 (ctrl 2) contains no scavengase p20. Results are representative of duplicate determinations of duplicate preparations of purified scavengase p20. The standard deviation of each point is below 10%.

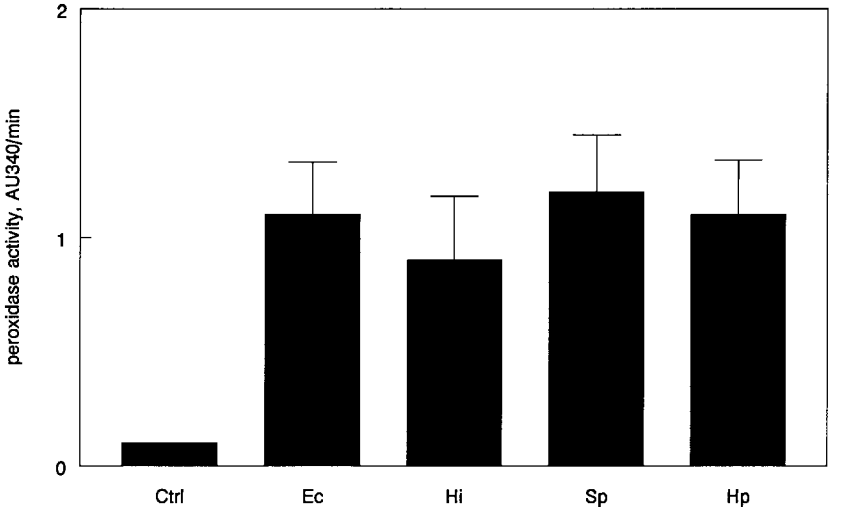

Fig. 5. Thioredoxin-linked peroxidase activity of purified scavengase p20 from *E. coli* (Ec), *H. influenzae* (Hi), *S. pneumoniae* (Sp) and *H. pylori* (Hp). No scavengase p20 was added to the control (Ctrl). Results are representative of triplicate experiments. Error bars indicate the standard error of the mean.

influenzae, *S. pneumoniae* and *H. pylori* all exhibit this activity to a level comparable to the previously characterized *E. coli* enzyme (Fig. 4). Furthermore, we showed that all these preparations of purified scavengase p20 possess a peroxidase activity dependent on the thioredoxin system (Fig. 5). The GS protection and the peroxidase activities were completely lost if the preparations were preincubated with a rabbit polyclonal antibody raised against *E. coli* scavengase p20 (data not shown).

Results from both assays corroborate the notion that scavengase p20 proteins from different species share common features not only in the amino acid sequence, but also in their enzymatic activities. These enzymes represent the prototype of a novel antioxidant family. Proteins in this family are currently found only in bacteria and they all possess a thioredoxin-dependent peroxidase activity. The mechanisms by which they scavenge H_2O_2 would be similar to those proposed for peroxiredoxins [7,9]. The cystine disulfides in AhpC has been suggested to be involved in the catalysis of peroxide reduction [7]. In this regard, the highly conserved cysteine residue in scavengase p20 (as shown in Fig. 1) would also play a crucial role. One member of the peroxiredoxin family was recently found to bind heme with very high affinity [24]. It would be of interest to see whether scavengase p20 would also bind heme. AhpC, thioredoxin and thioredoxin reductase have already been characterized in some of the bacteria described here [9,18]. Further studies are required to elucidate the interplay of these closely linked bacterial antioxidants and their effects on the host immune response.

3.3. Overlapping scavengase p20 and SOD genes in H. pylori

The functional implications of scavengase p20 in the biology and pathogenesis of *H. pylori* merit further discussion. The causative role of *H. pylori* in duodenal ulcer, chronic gastritis and gastric carcinoma is established [2,25] and the release of ROS from mucosal neutrophils is believed to play an important role in the pathogenesis [2,3]. In this scenario, the *H. pylori* antioxidant enzymes would be critically involved in the dysregulation of the host response.

It is very interesting that the *H. pylori* scavengase p20 gene is adjacent to the SOD gene but the polarity of transcription is opposite. The close proximity implicates that they might be members of an operon/gene cluster coregulated by the same transcription factor(s). We note that the DNA fragment encoding SOD and scavengase p20 was isolated by a functional complementation approach. SOD-negative *E. coli* transformed with this *H. pylori* DNA fragment was selected for survival on minimal agar supplemented with paraquat under aerobic conditions [6]. We postulate that in addition to the SOD gene the newly identified scavengase p20 gene in this fragment would also contribute to the survival.

The protein microsequencing data shown earlier in this study are suggestive of a sequencing error in the published sequence of this *H. pylori* DNA fragment. Taking this into account, the initiating ATG in the *H. pylori* scavengase p20 gene would be located on the antisense strand complementary to a region immediately between the putative −35 and −10 signals of the SOD gene (Fig. 6). In this context, it would be of great interest to see the possible antisense inhibition of the SOD transcription by the *H. pylori* scavengase p20 mRNA. We speculate that this would serve as a negative feedback control for superoxide dismutation. In this instance, the exact boundaries of the SOD and scavengase p20 transcripts and

```
                                                    ⇐ scavengase
  G  L  Q  Y  T  E  E  K  F
                              T  V  K  Q  M          SD
GAAGATTAACTATACAAAGAAGAAATTTCATTGAAAAACGTACCAATTTAGGAAA
CTTCTAATTGATATGTTTCTTCTTTAAAGTAACTTTTTGCATGGTTAAATCCTTTCTAATTGA
    -35                                                  -10
ATTGCTATATTGCGTTAGAATAATAGTCAATTAAAATTACTTTTAAAAATTTAAGAGATTGTT
AAAGTCGTTTCATTTTAAAAACCCCTTAAAATCTACAAAATTTGCATAAAGATCCATTTTTAA
GAAAAGATTTACCAAAAAGTATTAAAAAATGATTACAATACGGCTATCTGATCACAAGGAGAA
                                                        SD
     M  F  T  L  R  E  L  P  F  A  K  D  S  M  G
AAC ATG TTT ACA TTA CGA GAG TTG CCT TTT GCT AAA GAC AGC ATG GGA
    ⇒ SOD
```

Fig. 6. Nucleotide sequence of the junction of *H. pylori* SOD and scavengase p20 genes. The sequence was determined and the promoter elements for SOD were predicted by Spiegelhalder et al. [6]. Putative −10 and −35 signals, the Shine-Dalgarno (SD) sequence and the initiating ATG for the SOD gene are underlined. The putative SD sequence and the possible initiating ATG for the scavengase p20 gene are double-underlined. Arrows indicate the directions of transcription/translation. The putative amino acid sequence for the scavengase p20 is predicted from two reading frames and is thus presented in two separate lines.

promoters in *H. pylori* require further clarification and experimental analysis. Efforts are currently ongoing to reclone these genes from *H. pylori* and to decipher whether the expression of scavengase p20 would indeed inhibit SOD induction.

Taken together, the close vicinity of the two genes, a possible interaction of their transcripts and a functional linkage of these antioxidants may represent three levels (DNA, RNA and protein) of control mechanisms operative in the *H. pylori* antioxidant gene expression. Elucidation of these complex mechanisms would provide novel insights into the biology of this and other pathogenic bacteria.

Acknowledgements: The work was supported in part by the Chinese Academy of Preventive Medicine.

References

[1] V.L. Frazon, J. Arondel, P.J. Sansonetti, Infect Immun 58 (1990) 529–535.
[2] A. Lee, J. Fox, S. Hazell, Infect Immun 61 (1993) 1601–1610.
[3] A. Norgaard, L.P. Anderson, L. Elsborg, S. Holck, H. Nielsen, J Infect Dis 174 (1996) 544–551.
[4] E.C. Pesci, D.L. Cottle, C.L. Pickett, Infect Immun 62 (1994) 2687–2694.
[5] E.C. Pesci, C.L. Pickett, Gene 143 (1994) 111–116.
[6] C. Spiegelhalder, B. Gerstenecker, A. Kersten, E. Schiltz, M. Kist, Infect Immun 61 (1993) 5315–5325.
[7] L.B. Poole, Biochemistry 35 (1996) 65–75.
[8] L.B. Poole, H.R. Ellis, Biochemistry 35 (1996) 56–64.
[9] H.Z. Chae, K. Robison, L.B. Poole, G. Church, G. Storz, S.G. Rhee, Proc Natl Acad Sci USA 91 (1994) 7017–7021.
[10] M.-K. Cha, H.K. Kim, I.-H. Kim, J Biol Chem 270 (1995) 28635–28641.
[11] Y.-S. Lim, M.-K. Cha, H.-K. Kim, I.H. Kim, Gene 140 (1994) 279–284.
[12] L.E.S. Netto, H.Z. Chae, S.-W. Kang, S.G. Rhee, E.R. Stadtman, J Biol Chem 271 (1996) 15315–15321.
[13] S.F. Altschul, M.S. Boguski, W. Gish, J.C. Wootton, Nature Genet 6 (1994) 119–129.
[14] D.F. Feng, R.F. Doolittle, J Mol Evol 25 (1987) 351–360.
[15] J. Felsenstein, Methods Enzymol 266 (1996) 418–427.
[16] D.-Y. Jin, R.-H. Xu, Y. Zhou, P. Wang, Y.-D. Hou, Sci China (Sci Sin) Ser B 36 (1993) 1224–1232.
[17] H.Z. Chae, S.J. Chung, S.G. Rhee, J Biol Chem 269 (1994) 27670–27678.
[18] R.D. Fleischmann, M.D. Adams, O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult, et al. Science 269 (1995) 496–512.
[19] K.J. Hughes, K.D. Everiss, C.W. Harkey, K.M. Peterson, Gene 148 (1994) 97–100.
[20] J.S. Sampson, S.P. O'Connor, A.R. Stinson, J.A. Tharpe, H. Russell, Infect Immun 62 (1994) 319–324.
[21] J.C. Fenno, D.J. LeBlanc, P.M. Fives-Taylor, Infect Immun 57 (1989) 3527–3533.
[22] N. Ganeshkumar, P.M. Hannam, P.E. Kolenbrander, B.C. Mcbride, Infect Immun 59 (1991) 1093–1099.
[23] P.E. Kolenbrander, R.N. Andersen, N. Ganeshkumar, Infect Immun 62 (1994) 4469–4480.
[24] S. Iwahara, H. Satoh, D.-X. Song, J. Webb, A.L. Burlingame, Y. Nagae, U. Muller-Eberhard, Biochemistry 34 (1995) 13398–13406.
[25] J. Parsonnet, G.D. Friedman, D.P. Vandersteen, Y. Chang, J.H. Vogelman, N. Orentreich, R.K. Sibley, New Engl J Med 325 (1991) 1127–1133.