## Short sequence-paper

# Nucleotide sequence of the $v m h A$ gene encoding hemolysin from Vibrio mimicus 

Gu-Taek Kim ${ }^{\text {a }}$, Jong-Young Lee ${ }^{\text {a }}$, Sung-Hoi Huh ${ }^{\text {b }}$, Ju-Hyun Yu ${ }^{\text {c }}$, In-Soo Kong ${ }^{\text {a, }}{ }^{\text {x }}$<br>${ }^{a}$ RCOID and Department of Biotechnology and Bioengineering, Pusan 608-737, South Korea<br>${ }^{\text {b }}$ Department of Oceanography, Pukyong National University, Pusan 608-737, South Korea<br>${ }^{\text {c }}$ Bioproducts Research Center, Yonsei University, 134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, South Korea

Received 15 October 1996; revised 17 January 1997; accepted 20 January 1997


#### Abstract

The structural gene (vmhA) of hemolysin from Vibrio mimicus (ATCC33653) was cloned and sequenced. The $v m h A$ gene contains an open reading frame consisting of 2232 nucleotides which can code for a protein of 744 amino acids with a predicted molecular mass of 83059 . The similarity of amino acid sequence shows $81.6 \%$ identity with Vibrio cholerae El Tor hemolysin.


Keywords: Vibrio; Hemolysis; Enteropathogenic bacterium; Vm-hemolysin
V. mimicus is an enteropathogenic bacterium which inhabits aquatic environments and apparently causes diarrhea, usually after the consumption of uncooked seafood [1]. Several pathogenic factors of V. mimicus, including cholera toxin (CT) [2], CT-related enterotoxin [3], Escherichia coli heat-stable enterotoxin (ST)-like toxins [4,5] or protease [6], have been reported. Many $V$. mimicus strains isolated from environment are capable of causing diarrhea, even though they cannot produce these enterotoxins. Therefore, it was postulated that another toxin is involved in the bloody diarrhea which is one particular clinical symptom of V. mimicus gastroenteritis. In addition to the above toxins, hemolysins are suspected to be the pathogenic factor of the vibrio. Two kinds of hemolysin produced by $V$. mimicus, Vm-he-

[^0]molysin ( $M_{\mathrm{r}} 58000$ ), Vm-rTDH ( $M_{\mathrm{r}} 22000$ ), have been reported [7]. The former is immunologically cross-reactive with $V$. cholerae El Tor hemolysin and the latter is cross-reactive with V. parahaemolyticus thermostable direct hemolysin (TDH). The nucleotide sequence comparison of Vm-rTDH and Vp-TDH revealed that they were very homologous and had only minor variations but the flanking sequences of the hemolysin genes were dissimilar, indicating that they have a common ancestor and suggesting that they may have been transferred between vibrio species as a descrete genetic unit [8]. In this communication, we report the nucleotide sequence of gene encoding hemolysin similar to $V$. cholerae El Tor hemolysin.

A gene bank of the $V$. mimicus (ATCC33653) chromosomal DNA (partially digested Pst I/Sal I fragments) was prepared in the pUC19. One clone, pVMH 194 , was isolated by $\beta$-hemolysis on TSAII medium containing $5 \%$ sheep blood and the nucleotide sequence, containing an ORF of 2232 bp ,
was determined. Fig. 1 shows the nucleotide sequence and deduced amino acid sequence of the hemolysin, consisting of 744 aa with a predicted molecular mass of 83 kDa . Typical GAGGT, RBS or SD , sequence exists in 5 bp upstream from the ATG hemolysin initiation codon. However, the protein size $(83 \mathrm{kDa})$ is larger than that of previously purified

Vm-hemolysin ( 58 kDa ) [7]. In $V$. cholerae El Tor hemolysin, the 82 kDa preprotoxin synthesized in the cytoplasm is secreted through the membrane into the culture medium as the 79 kDa inactive protoxin after cleavage of the signal peptide and is then further processed into the 65 kDa active hemolysin by release of the N-terminal 15 kDa fragment [9]. Thus,
1 CTGCAGATCTGGTCAGATCTCTCTTGTATTGTGAATGAGACATTTGTCTGGTCAATGACAAGTGCTCTGTCATTACTCAGGTTGGAAGTAAGATCAAOCTTGCGTATAAAGTCTCCTATT
$P_{5} t \mathrm{I}$
121 ATTCTGTTCATAATTGATCTTTATTTTTCATATTAATGGACTTAAGTGATATTTAATGATTGAAAGTTAATCACTTAATGAATTTCTATTTTTATCTATTTAATTCAATCTAATTTAGIT
241 CAAATTAAATTAGTGCGAATGAATAATGTAAATATCCAAAAAACAAATTGGGAATTAATAGTTGTGTGAAGAATTAATTCTCGCCAAATAATAAGTTTTCAAGCGTGAATTTGTGATCTC
361 сTGTTACTATTTTCTGTTACAACTATTTACTTTTGGCATGTTTAGGGATATTCCAATAAGTGTGGTGGCGGAGAATATAAACGTATTCTTTCAAATAAAAGAGGTGGAATATTCTGTCAA
481 AGGGGTATGTGTTAAGGCGTAGTTCTGCCCCAAGCCACACACAACATCAAGGATGACGACGGTAACCCACGAGAAGCATTCAAACTGGGTTTGCTEGATGATTTGAACCTGCTGTGGTGT
601 TGATTAAATTTATCCTCTTAAAGCATAACGCTTAATTATGTGCGTAGTGATGAATGAATTTCCCGCTAGTCTGTTCTTTCAGTGGTTCGAGTTAATTCACTTTGTTTGAAAATCTCTCTT
721 СTAATAACACTAAAAATAACTGAATCAGTGAGGTTTATATGCCAAAACTCAATCGTTGTGCAATCGCAATCCTAACGATATTCAGOGCAATATCTAGTCATACTGTGTTOGCGAATATCA
S.D. M P K L N R C A I A I L T I F S A I S S H TVFAN I 27
841 GTGATCCOGTTGGTGAGGCTGTTGAAATTATTAGCCAAGTCGCGGATAACCACGCAATAAAATATTACAACGCTGCAGATTGGCAATCAGAAGAAAGCCAATTACCCAGTTTAGCTGAGT
S D PVGE A VE I I S Q V A D N H A I K Y Y N A A D Q S E E S Q L P S L A E 67
961 TGCGTGAGCAGGTGATTAACCAGCAGTTGCGTATCTTGGTGGATTTCAGCCAGATCTCTGACCCTGATCGCCTAACCGAAATGCAGGTGAAGTTTAGAAAAACCTATGGTGTTGGGTTG
LREQVINQ QLRILVDFSQIS DPDRLTEMQVKFRKTYGVGF107
1081 CTAATACATTCATCGTAATCACTGAGCATAAAGGTGAGCTGCTGTTTACACCGTTTGATCGTGCGGAGGATGTGGATCCAACACTGCTGGAGGCCOCACGTTCTGAGCGTGTGTTGAGTC
A V T F I V I T E H K G E L L F T P F D R A E D V D P T L L E A P R S E R V L S 147
1201 GGTCACGTCGCTCAGTGTCTGCCAATAATGTCACTAACAACAATGAAACGAACACGTTGCCGCACGTTGCTTTTTACATCAACGTAAACOGTTCAATCAGCGATGAAGAGTGTACTTTCG
R S R R S V S A N N V T N N N E T N T L P H V A F Y I N V N R S I S D E E C T F 187
1321 CCAATTCTTGGCTATGGAAAAATGACAAAGGCAGTCGTCCATTTTGTAAAGATGCCAATATCTCATTGATTTACCGTGTGAACTTAGAGOGTTCATTGCAGTATGGCATTGTGGGTTCGG
A N S L L K N D K G S R F F C K D A N I S L I Y R V N L E R S L Q Y G I V G S 227
1441 CAACTCCGGATGCCAAAATTGTGGGTATCAGCTTAGATGACGACAGCACTGGCGCGGGTATCCATCTCAATGATCAACTTGGTTATCGCTTTTTCAAAGCAGGATATACGACGCTOGATG
A T P D A K I Y R I S L D D D S T G A G I H L N D Q L G Y R F F K A G Y T T L D 267
1561 CTTACTTCCGTGAGTGGTCAACGGATGCGATAGCGCAGGATTATCGCTTCAGTTTTAACGCTTCGAATGATAAGGOGCAGATTTTAAGGACTTTCCCGGTCACTAACGTCAACGCAAACT
A Y F R E S T D A I A Q D Y R F S F N A S N D K A Q i L K T F P V T N V N A N 307
1681 TTGAGCGTAAAGAAGTCTCTGGTTTTGAGCTTGGTGTAACGGGCAGTGTTGAAGCGGATAAAAATGGCCCGHAGGTAAGTTGGAAGCGAAAGCCAGCTACACGCAAAGCCGTGGTTGA
F E R K E V S G F E L G V T G S VEAD K N G P K A K L E A K A S Y T Q S R L 347
1801 CCTACAACACGCAGGATTATCGAATTGAGCGGAGTGCGAAAATGCGCAAASGITAGCTTTACTTGGACCGOCAGCAGTACGCAACGGCTGAATCTCTGCTCAACCGTTCAACGGATG
T Y N T Q D Y R I E R S A K N A Q N V S F T N R Q Q Y A T A E S L L N R S T D 387
1921 CATTGTGGGTGGAACGTATCCTGTTGATGTGAAOCGCATCAGCCCGCTGAGTTACGCCAGTTTTGTGCCAAAAATGGATGTAATTTACAAAGCGTCACCGAAAGAGACGGGGGGACCG
A L Y V E T Y P V D Y N R I S P L S Y A S F Y P K M D Y I Y K A S P K E T G S T 427

2041 TTTTTGTCATCGACTCTTCGGTCAATATCCGTCCTATCTATAATGGTGCTTACAAACACTATTATGTGGTGGGTAGTCATCAGTCTTATCACGGGTTTGAAGATACGCCTCGTCGTCGTG
 2161 TGACAAAATCGGCAAGCTTTACTGTGGATTGGGATCAOCCTGTATTTACTGGTGGTCGTCCAGTGAATTTACAGCTGGCGAGCTTCAACAATCGCTGTATTGAAGCTGATGATCAAGGTC
 2281 GCTTGATGGCAACTACCTGTGACAGTCAGCAAGCGGCGCAATCTTITATCTATGATCAACAAGGTCGTTACGTGAGTGCAAGTAACACCAAGCTTTGTCTGGATGGTGAGGCGCTGGATT R L M A T T C D S Q Q A A Q S F I Y D Q Q G R Y Y S A S N T K L C L D G E A L D 547 2401 CTCTGCACACCTGTAACCAGAACTTGACCCAACGTTGGGAGTGGCGTGAAGGTTCTGACGAACTGAGCAATGTGTTTAAOGGTGAAGTATTGGGACATGACAAACAAACCGGGGAACTCG
 2521 GOCTTTATTCAACAGGCAGTGATGCAGTGAGCCTACGTACGATTACGTCTTACACCAATGITTTCCATGAACAGGAAAGCTCACCCGTGCTTGGTTTAACCCAAGGGAAAGTGAATCAGC G L Y S T G S D A V S L R T I T S Y T N V F H E Q E S S P V L G L T Q G K V N Q 627 2641 AGCGAGTAGAAAAGGATAACCAACTCTACGTACGTGCAGGTGCOGCCATTGATGCTTTAGGAACTGCTCCTGAGTTGTTGGTTGGTGGTACGGGCGGTAGCATCACAACGGITGATCTCA Q R V E K D N Q L Y V R A G A A I D A L G T A P E L L V G G T G G S I T T V D L. 667 2761 CTGGGCTTAGTTCTATCACAGCGACTTCTGGCGACTTCAACTTCGGTGGCCAGCAGTTGGTGGCATTAACCTTTACTTACCAAGATGGGCGTCAACAAATGGTCGGTTCGAAAGAGCACA
 2881 TCAGCAATGCTCATGAGGAOGCTTTGACGTAOCGACAGGAGCCAATATTACCCAAATGAAAGTTTGGGCAGATAACTGGTTGGTGAAAGGTGTGCAGTTTGACCTGAACTAAAATGTGA
 3001 TTAATGCCCTGATAATTCAGGGCATTAATGTTCATAATACTTTTAATGTTTCTTATTGACTGATTTGTCGAC SalI

Fig. 1. The complete nt and deduced aa sequence of the $v m h A$ gene. The nt sequence is numbered on the left of the sequence and the deduced aa sequence on the right. Asterisk represents the stop codon. Rho-independent stop region of mRNA is indicated by arrows. The GenBank accession number is U68271.


Fig. 2. Alignment of predicted amino acid sequences of $V m h A$ from V. mimicus and HlyA from V. cholerae O1 Biotype El Tor. Identical amino acid residues in the two hemolysins are boxed. Conserved cysteine residues are marked by asterisks.
we assume that Vm-hemolysin may also have very high possibility of the same two-step processing.

The hemolysin shared $81.6 \%$ identity with $V$. cholerae El Tor hemolysin, consisting of 741 aa with a predicted molecular mass of 82 kDa , having only a major variation of three amino acid deletion from 148 to 150 [9-11] (Fig. 2). However, no sequence homology with other hemolysins/cytolysins was found. The placements and numbers of cystine residues for Vm -hemolysin and El Tor hemolysin were identical, reflecting the similarity of their secondary structures. According to the comparison of cysteine residues, Vm-hemolysin might be divided into three regions; the N-terminal region (Met-1 to Cys-185), the central region (Cys-185 to Cys-552) and the C-terminal region (Cys-552 to Asn-744). The N-terminal region, showing relatively lower homology and having the variation with El Tor hemolysin, seemed to be less important for hemolytic activity. It was supported by
the fact that $V$. cholerae El Tor hemolysin is processed twice at Asn-26 and Asn-158, and the mature hemolysin has higher activity than the precursors by removing the N -terminal region [9]. The central region (Cys-185 to Cys-552) with six cystines had higher homology than other regions. The suggestion had been reported that the C-terminal region of El Tor hemolysin may be involved in the proper configuration of the protein for maximal hemolysin activity [12,13]. Thus, it might have an important role for the hemolysin activity although the C-terminal region of Vm-hemolysin has less homology than the central region. In addition, it will be required to explain the biochemical differences between Vm-hemolysin and El Tor hemolysin. The function of the central and C-terminal region of Vm-hemolysin will be elucidated by further study.

This research was supported in part by a grant from Engineering Research Center and by Non Directed Research Fund, Korea Research Foundation.

## References

[1] Shandera, W.X., Johnston, S.M., Davis, B.R. and Blake, P.A. (1983) Ann. Intern. Med. 99, 169-171.
[2] Chowdhury, M.A.R., Aziz, K.M.S., Kay, B.A. and Rahim, Z. (1987) J. Clin. Microbiol. 25, 2200-2203.
[3] Davis, B.R., Fanning, G.R., Madden, J.M., Steigerwalt, A.G., Bradford, H., Jr., Smith, H.L. and Brenner, D.J. (1981) J. Clin. Microbiol. 14, 631-639.
[4] Gyobu, Y., Kodama, H. and Uetake, H. (1988) Microbiol. Immunol. 32, 565-577.
[5] Nishibuchi, M. and Seidler, R.J. (1983) Appl. Environ. Microbiol. 45, 228-231.
[6] Chowdhury, M.A.R., Miyoshi, S. and Shinoda, S. (1991) Biomed. Lett. 44, 31-34.
[7] Honda, T., Narita, I., Yoh, M. and Miwatani, T. (1987) Jpn. J. Bacteriol. 42, 201.
[8] Terai, A., Shirai, H., Yoshida, O., Takeda, Y. and Nishibuchi, M. (1990) FEMS Microbiol. Lett. 71, 319-324.
[9] Yamamoto, K., Ichinose, Y., Shinagawa, H., Makino, K., Nakata, A., Iwanaga, M., Honda, T. and Miwatani, T. (1990) Infect. Immun. 58, 4106-4116.
[10] Manning, P.A., Brown, M.H. and Heusenroeder, M.W. (1984) Gene 31, 225-231.
[11] Rader, A.E. and Murphy, J.R. (1988) Infect. Immun. 56, 1414-1419.
[12] Alm, R.A., Mayrhofer, G., Kotlarski, I. and Manning, P.A. (1991) Vaccine 9, 588-594.
[13] Honma, Y., Yamamoto, K. and Iwanaga, M. (1995) FEMS Microbiol. Lett. 133, 151-154.


[^0]:    ${ }^{*}$ Corresponding author. Fax: +8251 6206180; E-mail: iskong@dolphin.pknu.ac.kr

