Biochimica et Biophysica Acta 1800 (2010) 629–638

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Activity of recombinant cysteine-rich domain proteins derived from the membrane-bound MUC17/Muc3 family mucins

Samuel B. Ho [a,b,*], Ying Luu [a,b], Laurie L. Shekels [c], Surinder K. Batra [d], Brandon Kandarian [b], David B. Evans [e], Phillip G. Zaworski [e], Cindy L. Wolfe [e], Robert L. Heinrikson [e]

[a] Department of Medicine, University of California-San Diego, San Diego, CA, USA
[b] VA San Diego Healthcare System, San Diego, CA, USA
[c] Research Service and VA Medical Center and University of Minnesota, Minneapolis, MN, USA
[d] Department of Biochemistry, University of Nebraska, Omaha, NE, USA
[e] Proteos, Inc., Kalamazoo, MI, USA

ARTICLE INFO

Article history:
Received 11 July 2009
Received in revised form 12 March 2010
Accepted 15 March 2010
Available online 20 March 2010

Keywords:
Mucin
Cell migration
Inflammatory bowel disease
MUC17
Epidermal growth factor
Recombinant protein

ABSTRACT

Background: The membrane-bound mucins, MUC17 (human) and Muc3 (mouse), are highly expressed on the apical surface of intestinal epithelia and have cytoprotective properties. Their extracellular regions contain two EGF-like **C**ys-**r**ich **d**omains (CRD1 and CRD2) connected by an intervening linker segment with SEA module (L), and may function to stimulate intestinal cell restitution. The purpose of this study was to determine the effect of size, recombinant host source, and external tags on mucin CRD1-L-CRD2 protein activity.

Methods: Four recombinant Muc3-CRD proteins and three MUC17-CRD proteins were generated using *Escherichia coli* or baculovirus-insect cell systems and tested in colonic cell cultures for activity related to cell migration and apoptosis.

Results: N-terminal glutathione-S-transferase (GST) or C-terminal His_8 tags had no effect on either the cell migration or anti-apoptosis activity of Muc3-CRD1-L-CRD2. His-tagged Muc3-CRD1-L-CRD2 proteins with truncated linker regions, or the linker region alone, did not demonstrate biologic activity. The human recombinant MUC17-CRD1-L-CRD2-His_8 was shown to have anti-apoptotic and pro-migratory activity, but did not stimulate cell proliferation. This protein showed similar in vitro biologic activity, whether produced in *E. coli* or a baculovirus-insect cell system.

Conclusions: Recombinant mucin proteins containing a bivalent display of Cys-rich domains accelerate colon cell migration and inhibit apoptosis, require a full-length intervening Linker-SEA segment for optimal biologic activity, and are functional when synthesized in either *E. coli* and insect cell systems.

General Significance: These results indicate that an *Escherichia coli*-derived full-length His_8-tagged human MUC17 CRD1-L-CRD2 recombinant protein is a biologically active candidate for further development as a therapeutic agent.

Published by Elsevier B.V.

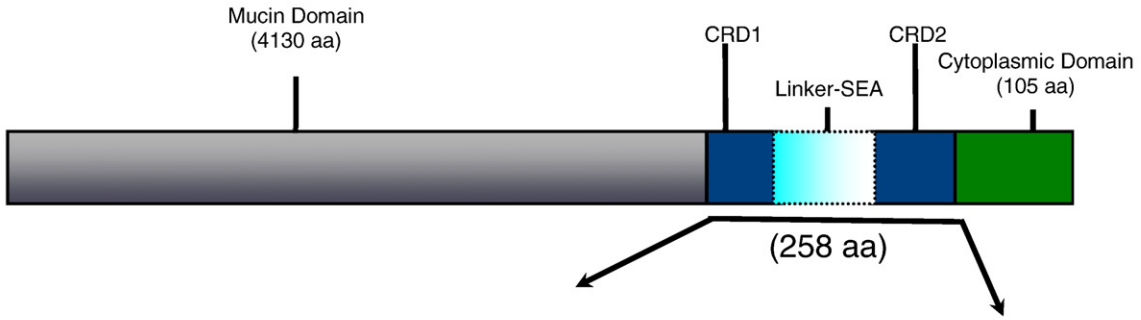
1. Introduction

Mucin-type proteins, categorized as secretory or membrane bound (1–3), represent the major structural proteins of mucous gels that are integral to the epithelial defense of respiratory, digestive, ocular, and reproductive surfaces. The membrane-bound mucins are characterized by an extracellular region with a short amino terminal domain, followed by a large, heavily O-glycosylated tandem repeat domain which accounts in part for their protective function. This large (>4000 amino acids) structural domain is followed by a globular region just proximal to the membrane that is made up of two Cys-rich motifs (CRD1 and CRD2), each with similarity to epidermal growth factor (EGF), separated by a Linker-SEA (L) domain which contains an SEA (sea urchin sperm protein, enterokinase, and agrin) module and Linker residues before and after the module. SEA domains are found in several membrane associated mucins and other membrane proteins and are thought to be important in the noncovalent association of protein subunits, and may play a role in the release of membrane protein subunits at the cell surface [1,2]. C-terminal to the CRD1-L-CRD2 unit is a transmembrane segment followed by a small cytoplasmic domain. An overview of this structure is shown schematically in Fig. 1A, together with a breakdown of amino acid sequences in each part of the CRD1-L-CRD2 segment of human MUC17 [3]. The related membrane bound mucin genes *MUC3A/B,*

* Corresponding author. Gastroenterology Section 111-D, VA San Diego Healthcare System, 3350 La Jolla Village Drive, San Diego, CA, USA. Tel.: +858 552 8585x4 2631; fax: +858 552 4327.
E-mail address: samuel.ho2@va.gov (S.B. Ho).

0304-4165/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.bbagen.2010.03.010

A Schematic layout of human MUC17 and sequences of CRD1, the Linker-SEA region (L) and CRD2



CRD1: RTTTCFGDGCQNTASRCKNGGTWDGLKCQCPNLYYGELCEEVVS
LINKER-SEA (L): SIDIGPPETISAQMELTVTVTSVKFTEELKNHSSQEFQEFKQTFTEQMNIVYSGI
PEYVGVNITKLRLG*SVVVEHDVLLRTKYTPEYKTVLDNATEVVKEKITKVTTQQI
LINKER (Lcys): MINDICSDMMCFNTTGTQVQNITVTQYDPEEDCRKMAKEYGDYFVVEYRDQKPYCIS
CRD2: PCEPGFSVSKNCNLGKCQMSLSGPQCLCVTTETHWYSGETCNQGTQK
Underlined sequence = SEA module; *=SEA cleavage site sequence

B Comparisons of human EGF and the CRD1 and CRD2 Domains in MUC Proteins

```
S-S Pairings        C6----------------C20
In EGF                    C14----------------C31
                                  C33----------C42

Human EGF      NSDSECPLSHDG--YCLHDGV---CMYIEALDKYACNCVVGYIGER----CQYRDLKWWELR
MUC17 CRD1      PRTTTCFG--DG---CQNTASR--CKNGGTWDGLKCQCPNLYYGEL----CEEVVSSIDIGPPE
Muc3  CRD1     GDKCICPNGFGGD-RCENIVNVVNCENGGTWDGLKCQCTSLFYGPR----CEELVE
MUC3  CRD1     QGQCACLPGFSGD-RCQLQT-R--CQNGGQWDGLKCQCPSTFYGSS----CEFAVE

MUC17 CRD2     YCISPCEPGFSVSKNCNLGK----CQMSLSGPQ--CLCVTTETHWYSGETCNQGTQKSL
Muc3  CRD2     FCITPCSAGYSTSKNCSYGK----CQLQRSGPQ--CLCLITDTHWYSGENCDWGIQK
MUC3  CRD2     RCVTKCTSGVDNAIDCHQGQ----CVLETSGPT--CRCYSTDTHWFSGPRCEVAV
```

C A New Structural Cys-containing Element in the Linker Segment, Lcys, Preceding CRD2

```
MUC17   MINDICS--DMMCFNTTGTQVQNITVTQYDPEEDCRKMAKE-YGDYFVVEYRDQKPYCISPC1  4 CYS
Muc3       NNC--SALLCFNSTATKVQNSATVSVNPEETCKKEAGEDFAKFVTLGQKGDKWFCITPC1  4 CYS
MUC3        SCQDSQTLCFKPDSIKVNNNSKTELTPAAICRRAAPTGYEEFYFPLVEATRLRCVTKC1  4 CYS
```

D Pre-CRD1 Sequences at the end of the Mucin-Rich Domain

```
MUC17          ...TTSFPTVTTTAVPTNTTIKSNPTSTPTVPRTTTC1FGD...  0 CYS
Muc3      ...TTTEVATTPEPTTTPAPTTTAVNCMNGGFWTGDKCIC1PNG...  2 CYS
MUC3                ...TSQMTTQSTLTTTAGTCDNGGTWEQGQCAC1LPG...  2 CYS
```

Fig. 1. (A) Schematic representation of MUC17 and amino acid sequences of units within the 259-residue Cys-rich region, including CRD1, the Linker-SEA (L) segment including the Lcys segment, and CRD2 (see references [2,3]). These sequences correspond to the core structure of MUC17-CRD1-L-CRD2 discussed in the text, and defined in Fig. 2. (B) Comparative sequence analysis of the Cys-rich regions of human MUC17 and MUC3, and mouse Muc3- comparisons of CRD units with themselves and with human EGF. (C) Comparisons within the segment of structure preceding CRD2, herein defined as Lcys. (D) Comparisons within the segment preceding CRD1. Note that these residues are seen in MUC3 and Muc3, but not MUC17; however, whether this results in differences in structure or aggregation of these proteins is not determined.

MUC12, and *MUC17* are clustered on chromosome 7q22, and are highly expressed in intestinal tissues at the apical surface of enterocytes. The mouse *Muc3* gene[4] is most similar in sequence and chromosomal localization to the human *MUC17* gene [3,5–7]. The similarity of Cys spacings in mouse Muc3 and human MUC3 and MUC17 mucin CRD proteins and EGF are indicated in Fig. 1B. The amino acid sequence of the remaining areas contain the 167 amino acid Linker-SEA domain, which contains a 106 amino acid SEA module, a 7 amino acid sequence preceding the SEA module, and a 54 amino acid region of unknown function that flanks the SEA module.