FEBS Letters 348 (1994) 109–113

FEBS 14183

Hypothesis

KEKE motifs

Proposed roles in protein–protein association and presentation of peptides by MHC Class I receptors

Claudio Realini[a], Scott W. Rogers[b], Martin Rechsteiner[a,*]

Departments of [a]Biochemistry and [b]Neurobiology and Anatomy/VA-GRECC, University of Utah, Salt Lake City, UT 84132, USA

Received 10 April 1994; revised version received 19 May 1994

Abstract

A stretch of 28 'alternating' lysine (K) and glutamate (E) residues is found in an activator of the multicatalytic protease. Such 'KEKE sequences' are also present in subunits of the multicatalytic protease, in subunits of the 26S protease and in a variety of chaperonins. We propose that KEKE regions promote association between protein complexes. Furthermore, they may contribute to the selection of peptides presented on MHC Class I receptors.

Key words: Multicatalytic protease/proteasome; Antigen presentation; KEKE motif; Chaperonin

1. Introduction

Two large proteases are found in the cytosolic and nuclear compartments of eukaryotic cells. One enzyme is known as the multicatalytic protease (MCP) or 20S proteasome [1–3]. MCP is composed of 24–28 subunits with a combined mass of about 650 kDa and it appears to be a hollow cylinder consisting of four stacked rings (see Fig. 1). MCP is the proteolytic core of an even larger enzyme, the 26S protease. In the presence of nucleotides, MCP cylinders associate with a regulatory ATPase complex (AC) that contains 15 or more polypeptides [4] ranging in molecular weight from 25 to 110 kDa (see Fig. 1). At least three ATPases [5] and a component that binds polymers of ubiquitin [6] are present among subunits of the regulatory complex. For certain substrates degradation by the 26S protease requires prior attachment of Ub chains [7]. However, recent studies demonstrate that unmodified ornithine decarboxylase is degraded by the 26S enzyme [8].

MCP also associates with a protein hexamer [9,10] that we call the regulator (REG). Association of MCP and REG does not require nucleotides and does not convert MCP to a proteinase. REG binding does, however, result in substantial activation of peptide hydrolysis. We recently cloned and expressed an intact cDNA encoding the human REG subunit [11]. Conceptual translation of the cDNA results in a protein 249 amino acids long. Near the center of the sequence, sandwiched between two regions characteristic of globular domains, there are 28 'alternating' lysine and glutamic acid residues (see Fig. 2). We call this unusual feature of the protein's sequence a 'KEKE motif', and it prompted the hypotheses presented here. We propose that KEKE motifs promote association between proteins. We further suggest that the regulator recognizes KEKE-like sequences in cellular or viral proteins and directs adjacent regions for presentation on MHC Class I receptors.

In reference 11, we defined KEKE motifs as greater than 12 amino acids in length, devoid of W, Y, F or P, more than 60% K and E/D, and lacking five positive or negatively charged residues in a row. KEKE motifs so defined are present in two α subunits of MCP, in subunit 12 of the ATPase complex that forms the 26S protease and in a number of chaperonins (Fig. 2). Of the more than 100,000 entries in PIR Library (release 39), 106 proteins were found to contain such KEKE motifs. Equivalent regions composed of arginine (R) and aspartate (D) are present in only two proteins, RD protein and hnRNP70 [12,13]. Considering all proteins in the data bank, the abundance of arginine is comparable to that of lysine; similarly aspartate and glutamate are present at equivalent frequency. Yet there are only two RDRD proteins and more than 100 KEKE proteins. Thus, KEKE motifs like those presented in Fig. 2 are not simply statistically expected arrangements of amino acids.

Our proposal that KEKE motifs promote protein:protein association is based in part on their presence in four proteins known to bind each other. It is well documented that MCP associates both with AC and REG [4,9,10,14]. Each of these multisubunit com-

*Corresponding author.

SSDI 0014-5793(94)00569-H

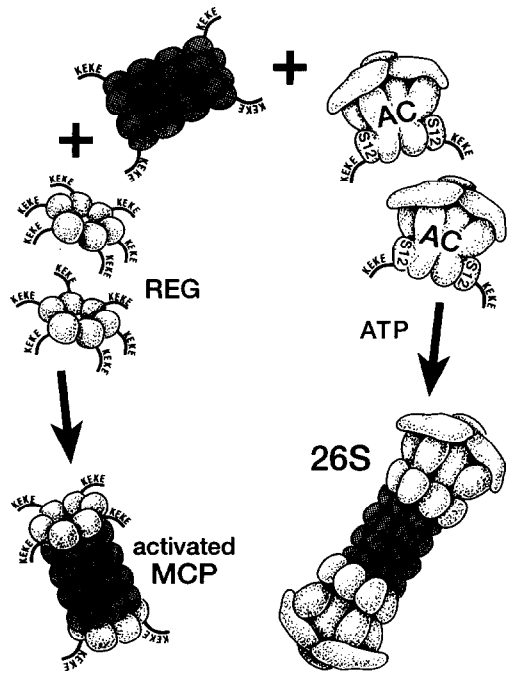

Fig. 1. Activation of the multicatalytic protease (proteasome). The multicatalytic protease is shown associating with two protein complexes, the ATPase complex (AC) and the regulator (REG), that activate its peptidase activities. We propose that KEKE regions mediate such associations.

plexes contains at least one component with a strong KEKE motif (Fig. 2). Since we have also found that MCP can form a stable complex with Hsp 90 (Hough and Rechsteiner, unpublished), four proteins with KEKE sequences bind to one another. Although such interactions may not involve the KEKE motifs per se, it seems reasonable to speculate that these regions are responsible for the observed associations. In fact, two studies have implicated KEKE motifs in protein–protein interaction. Nobel et al. [15] identify a KEKE region in MAP 1B as being responsible for its binding to microtubules. Likewise, Lu et al. [16] found that the interaction of dihydropyridine receptors with the calcium release channel involves KEKE regions.

Preliminary experiments in our lab also provide direct evidence for KEKE-mediated association of proteins. Using the KEKE motif from REG, we have produced a ubiquitin KEKE fusion peptide, (Ub-PVKEKE-KEERKKQQEKEDKDEKKKGED). This chimera binds the REG whereas Ub does not; moreover, Ub-KEKE potentiates the activity of the REG (Realini and Rechsteiner, in preparation). We do not know whether binding results from direct KEKE:KEKE interactions or association between the KEKE motif and a KEKE-binding site. However, future experiments involving model peptides should resolve this issue.

Our second hypothesis is an extension of the idea that KEKE sequences promote protein association. We propose that the regulator associates with KEKE-like sequences in cellular proteins or proteins from pathogens. We envision that this association is transient and that it results in MCP-mediated cleavage of the cellular/pathogen proteins. Like many others [17–19], we believe that the peptides generated by MCP are eventually transported to the cell surface for presentation to lymphocytes. Current views of antigen presentation by Class I receptors encoded in the major histocompatibility locus (MHC) invokes cytosolic proteolysis of cellular, viral or other parasitic proteins to produce peptides [17–19]. These peptides are transported into the endoplasmic reticulum (ER) lumen where they bind a groove in the MHC I receptor that accommodates protein fragments 8 to 10 amino acids in length. Association of the MHC I receptor with a tight-binding peptide and β_2 microglobulin releases it from calnexin, an 88 kDa chaperonin embedded in the ER membrane. The MHC I:peptide:β_2 microglobulin complex then travels to the cell surface. A number of questions concerning this pathway are unresolved, and some are controversial. What cellular protease(s) generates the peptides? How big are the peptides? How are they selected? Is peptide transport into the ER energy-dependent? Does peptide trimming occur in the ER?, etc.

Our hypothesis, which focuses on how peptides might be selected from precursor proteins, resulted from inspection of the JAK 1 kinase sequence [20]. Cells typically display about 50,000 MHC I receptors on their surface. Most peptides bound to MHC I receptors are present at 10 to 100 copies per cell. However, there are 10,000 copies of the peptide, SFFPEITHI, bound to MHC I receptors on p815 cells [18]. SFFPEITHI, which originates from JAK 1 kinase, is found in the sequence KEKEKNKLKRKKLENKDKKDEEKNKIREEWNN FSFFPEITHI. Given the proximity of such a highly presented peptide to a strong KEKE motif, we examined the sequence context of a large number of presented peptides (see Table 1). Inspection of these sequences revealed a high occurrence of KEKE-rich regions in precursors to presented peptides. Ten of the 54 sequences (18.5%) in Table 1 contain KEKE regions, whereas only 106 of the 100,346 entries in PIR LIB (release 39) contain

REGULATOR	....KEKEKEERKKQQEKEDKDEKKKGEDEDK.....
PROTEASOME C9	....KKHEEEEAKAEREKKEKEQREKDK .....
PROTEASOME 28.1	....KIIEKEKEEELEKKKQK .....
26S SUBUNIT 12	....EKKEGQEKEESKKDRKEDKEKDKDKEKSDVKKEEKK...
CALNEXIN	....KEEEEEKEEEKDKGDEEEEGEEKLEEKQKSDAE .....
JAK 1 KINASE	....EKEKNKLKRKKLENKDKKDEEKNKIREE .....
hHSP90	....EEKEDKEEEEKEKEEKESEDK .....
yHSP70	....EKLAAQRKAEAEKKEEKKDTE .....

Fig. 2. KEKE motifs in chaperonins, proteolytic components and a γ-IFN activated protein kinase. KEKE regions from the sequences of selected proteins are presented in the one letter amino acid code. Proteins were included on the basis of their likely involvement in antigen presentation.

stringent KEKEs. We estimate that there is a 5- to 10-fold redundancy in PIR LIB so the frequency of KEKEs lies between 106 in 20,000 and 106 in 10,000 distinct proteins. At most, KEKE-plus proteins constitute 1% of the proteins in PIR LIB. Thus, there is at minimum a 20-fold enrichment for strict KEKE motifs in proteins presently known to generate MHC I presented peptides. Another six proteins almost conform to our definition of a KEKE motif. Based on the sequences in Fig. 2 and Table 1, we developed two algorithms to identify KEKE regions (KEKE Searches). One algorithm identifies all the proteins in Fig. 2, 7 of 52 precursors to MHC I presented peptides and 3.7% of the 33,329 entries in the SWISS.PROT data bank (release 27). The other algorithm, which is less stringent, identifies 41% of the known precursors to MHC I presented peptides and 16% of the entries in SWISS.PROT. These and any newly developed algorithms will be available from the authors upon request.

The idea that KEKE motifs promote association can

Table 1
KEKE sequences in precursors to presented peptides

Strict KEKEs

[K E K E K N K L K R K K L E N K D K K D E E K N K I R E E] W N N F S F F P E I T H I	JAK1
[E E E K K K M E E S K A K] F E N L C K L M K E I L D K K V E K V T	HSP90
[K K E K A R V I T E E E K] N F K A F A S L R M A R A N A R L F G I R A K	BBC1
[E K L D I K L D S E D K D K E G K] P L L K A V M R R W L P A G D A	eEF2
[E T E D N K E K K S A K D] A L L L W C Q M K T A G Y P N V N I H N F	Spectrin
[K P A E K K D D L K E E K K D D L P K E E K K D D L P	
K E E K K D D P P K E E K K D D L P K E E K K D A P K D	
G N K D A P K E E K K A D P P K E] ...(VADPNAP)$_n$...E D S Y V P S A E Q I L E F V K Q M	P. yoelii
[E G K K N E K K N E K I E R N N K]-(pro rich)-N N D D S Y I P S A E K I L E F V K Q I	P. berghei
[K E G A D K E K K K E K G K E K E E E]-(GQAP)$_n$-N E K V V N D Y L L H K I R S S V T T	P. knowlsei
[K D T K E A L D K I E E E Q N K S K K K]-(Q-rich)-R W I I L G L N K	HIV gag
[K E I C E K N D E C E S S K E]-(62)-N I R T L I Q I L K Q K I A D L	IL6 precurs.
R R I K E I V K K-(13)-[E K E R D K E V S D D E A E K E D K E E E K E K E E K E S E D K]	HSP90
A R L F G I R A [K R A K E A A E Q D V E K K K]	BBC1

Close KEKEs (Almost 60% and/or one excluded residue)

[R A E E E D K K E D]-(60)-E N T V F D A K R L I G R	BIP
[E E E G K G K D A S G N K V K A E]-(34)-G V N L P Q K A G G F L M	PGK
[E M K K R E S K F I K D A D E E K]-	
(15)-[E K D A E L E K L R N E]-(24)-K V K L E L K V K N L E L	restin
E Y L N K I Q N S L S T E W S P C S V T-(17)-[K D E L D Y A N D I E K K I C K M E K]	P. falc.
Q G I N N L D N L-(~150)-[E D S Q E N A D K N E D G G E K]	T.Ag.polyoma
K Y Q A V T T T L - [E E K R K E K A K I H Y R]	T.A. P198

Negatives

S I I N F E K L	oval.	L S P F P F D L	OGDH
V S D I V G P D G L V Y	fibrillarin	S S E Q T F M Y Y	ODC (human)
F A P G N Y P A L	Sendai	R G Y V Y Q G L	VSV Nuc Prot
G I L G F V F T L	Flu matrix	E Y L N K I Q N S L-	
M I E P R T L Q Y	ribo. S16	S T E W S P C S V T	P. falc.
T Q H N R A L D L	P91A	R R S K E I T V R	p68
L L F G Y P V Y V	HTLV-1	G Y K D G N E Y I	listeriol.
P K Y V K Q N T L K L A	Flu. Hema. p.	S S I E F A R L	herpes vir
T L W V D P Y E V	TIS21	V P K L K V C A L	ribos. L18
E A D P T G H S Y V	MAG-1 Ag	G I L G F V F T L	Flu matrix
L Y Q N V G T Y V	Flu hemagglut.	S A I N N Y A Q K L	TAg polyoma
R R Y P D A V Y L	measles fus. pro.	C K G V N K E Y L	TAq polyoma
H P K Y K T E L	tristet.	T Y Q R T R A L V	Flu
E P K Y K T Q L	PPAS (yeast)	G R A F V T I G K	HIV-Gp160
A S N E N M E T M	Flu hemagglut.	F R Y N G L I H R	ribo. L28
S G P S N T P P E I	E1A 32 K	G R I D K P I L	ribo. L8
S T G N L I A P E Y G F K I S	Flu (A,JAP)	R Y L K N G K E T L-	
R P Q A S G V Y M	LMCV N.P.	Q R A	HLA CW-3
I Y A T V A G S	Flu hemagglut.		

Large letters denote sequences of presented peptides. The smaller letters are amino acid sequences flanking them. The sequences are taken mainly from Falk and Rötzschke [25], Rammensee et al. [26], Jardetsky et al. [27], Hunt et al. [28], Di Brino et al. [29], and Corr et al. [30].

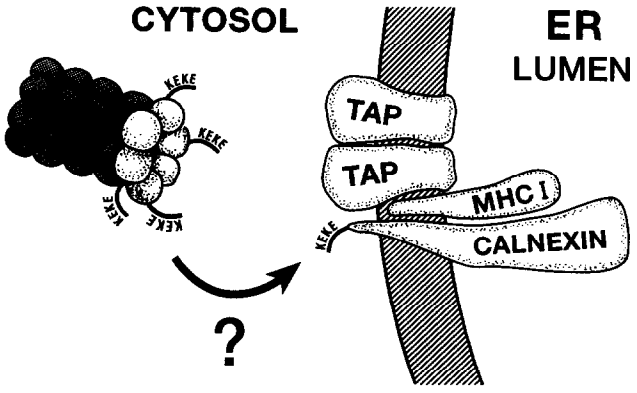

Fig. 3. Hypothetical KEKE-mediated association of regulator-MCP complex with calnexin. This diagram depicts our proposed mechanism for channeling peptides to the MHC I receptor. See text for a more complete description.

account for their significant enrichment in components of the 26S/MCP proteolytic pathway and for their enrichment in proteins that are the source of presented peptides. As noted above, KEKE motifs could, like leucine zippers, interact directly to mediate association. Alternatively, the REG may possess both KEKE sequences and KEKE binding sites. We consider both possibilities in the following discussion. The number of KEKE sequences (or KEKE-binding sites) in each component of the presentation pathway suggests how the process might occur. Two subunits located in the end rings of MCP contain C-terminal KEKE extensions. Each identical subunit of the REG hexamer possesses a KEKE motif (or KEKE-binding site). Thus, one can imagine that the REG binds MCP with either KEKEs or KEKE-binding sites to spare. The excess elements in the regulator would be available to bind KEKE-like regions in potential proteolytic substrates. A 40 residue KEKE sequence is present in the cytoplasmic tail of calnexin. As mentioned above, this chaperonin holds MHC I receptors in the endoplasmic reticulum until immunogenic peptides bind them. Given these observations, one can arrive at a reasonably detailed hypothesis for how cells generate peptides and transfer them to the MHC I receptor. REG binds substrates and MCP facilitating their cleavage by MCP. REG can also bind calnexin either by direct KEKE:KEKE interaction or by virtue of KEKE-binding sites. This results in peptides being generated adjacent to the ER membrane. Released peptides are transferred into the ER lumen, presumably by TAPs (transporters associated with antigen processing). These ideas are diagrammed in Fig. 3 except that substrates have been omitted for clarity.

By itself, REG does not enhance the cleavage of folded proteins, like lysozyme or bovine serum albumin [10]. One might, therefore, question our proposal that REG promotes precursor cleavage by MCP. It should be noted, however, that REG does enhance degradation of casein, and it promotes the degradation of certain iodinated proteins (see Fig. 12, panel B versus C in ref. 10). Moreover, it is possible that REG binds substrates when they are nascent chains or that REG collaborates with heat shock proteins; e.g. hsp 70 or hsp 90, to promote substrate degradation by MCP. Thus, REG may select substrates directly or in concert with additional proteolytic/chaperonin factors.

To reiterate, the presence of six KEKE regions (KEKE-binding sites) in the assembled REG raises the possibility that it can simultaneously associate with MCP, substrates and calnexin. By doing so, we imagine that the REG channels available MCPs into the antigen processing pathway. In fact, we have recently demonstrated that γ-IFN induces synthesis of the REG in HeLa cells [11] as has been observed for other postulated components of the antigen presentation pathway (e.g. TAP1/TAP2, MHC-Class I and Class II receptors, as well as the LMP2 and LMP7 subunits of the multicatalytic protease).

In summary, we have found significant enrichment of lysine-glutamate tracts in cellular components thought to be involved in antigen presentation and in precursors to presented peptides. Our hypothesis that these KEKE motifs are involved in protein association is directly supported by our studies with Ub-KEKE. However, it has not been demonstrated that KEKE regions are important for presentation of antigenic peptides. And we certainly do not suggest that all peptides presented on MHC I receptors are selected on the basis of KEKE regions. In fact, there is evidence that the 26S ATP/Ub-dependent protease can provide peptides for presentation by MHC I receptors [21,22]. It is interesting in this regard that 12 of 35 KEKE negative precursors in Table 1 contain PEST sequences [23] which are known to promote rapid proteolysis. And the 26S enzyme has been shown to degrade one PEST-plus protein, ornithine decarboxylase [8].

We have not postulated a structural basis for KEKE-mediated interactions. In this regard, however, we note that a similar motif, . . .$(EEEKKAAEERAKA)_{15}$ . . . present in caldesmon, is likely to be a helix [24]. Accordingly, we suspect that KEKEs form helices, and these, in turn, may form coiled-coils. Whatever the case, our hypothesis makes a number of readily testable predictions. And because the hypothesis has important implications for producing specific peptide:MHC I complexes, we believe it warrants experimental challenge.

Acknowledgements: These studies were supported by National Institutes of Health grants to Martin Rechsteiner (GM37009) and Scott Rogers (AG04418), by an award from the Lucille Markey Charitable Trust, and by a postdoctoral fellowship from the Swiss National Science Foundation to Claudio Realini. We thank Dana Carroll, Wolfgang Dubiel, Laura Hoffman and Vicenca Ustrell for helpful comments on the manuscript. We also appreciate the expert word processing of Linda Van Orden.

References

[1] Tanaka, K., Tamura, T., Yoshimura, T. and Ichihara, A. (1992) New Biol. 4, 173–187.
[2] Rechsteiner, M., Hoffman, L. and Dubiel, W. (1993) J. Biol. Chem. 268, 6065–6068.
[3] Rivett, J. (1993) Biochem. J. 291, 1–10.
[4] Hoffman, L., Pratt, G. and Rechsteiner, M. (1992) J. Biol. Chem. 267, 22362–22368.
[5] Dubiel, W., Ferrell, K. and Rechsteiner, M. (1994) Biol. Chem. Hoppe-Seyler's 375, in press.
[6] Deveraux, Q., Ustrell, V., Pickart, C. and Rechsteiner, M. (1994) J. Biol. Chem. 269, 7059–7061.
[7] Hershko, A. and Ciechanover, A. (1992) Annu. Rev. Biochem. 61, 761–807.
[8] Murakami, Y., Matsufuji, S., Kameji, T., Hayashi, S., Igarashi, K., Tamura, T., Tanaka, K. and Ichihara, A. (1992) Nature 360, 597–599.
[9] Ma, C.-P., Slaughter, C.A. and DeMartino, G.N. (1992) J. Biol. Chem. 267, 10515–10523.
[10] Dubiel, W., Pratt, G., Ferrell, K. and Rechsteiner, M. (1992) J. Biol. Chem. 267, 22369–22377.
[11] Realini, C., Dubiel, W., Pratt, G., Ferrell, K. and Rechsteiner, M. (1994) J. Biol. Chem., in press.
[12] Levi-Strauss, M., Carroll, M.C., Steinmetz, M. and Meo, T. (1988) Science 240, 201–207.
[13] Spritz, R.A., Strunk, K., Surowy, C.S., Hoch, O.S., Barton, D.E. and Francke, U. (1987) Nucleic Acid Res. 15, 10373–10391.
[14] Udvardy, A. (1993) J. Biol. Chem. 268, 9055–9062.
[15] Nobel, M., Lewis, S.A. and Cowan, N.J. (1989) J. Cell Biol. 109, 3367–3376.
[16] Lu, X., Le, X. and Meissner, G. (1994) J. Biol. Chem. 269, 6511–6516.
[17] Yewdell, J.W. and Bennink, J.R. (1992) Adv. Immunol. 52, 1–123.
[18] Rammensee, H.G., Falk, K. and Rötzschke, O. (1993) Annu. Rev. Immunol. 11, 213–244.
[19] Germain, R.N. (1994) Cell 76, 287–299.
[20] Wilks, A.F., Harpur, A.G., Kurban, R.R., Ralph, S.J., Zurcher, G. and Ziemiecki, A. (1991) Mol. Cell Biol. 11, 2057–1065.
[21] Townsend, A., Bastin, J., Could, K., Brownlee, G., Andrew, M., Coupar, B., Boyle, D., Chan, S. and Smith, G. (1988) J. Exp. Med. 168, 1211–1224.
[22] Michalek, M.T., Grant, E.P., Gramm, C., Goldberg, A.L. and Rock, K.L. (1993) Nature 363, 552–554.
[23] Rogers, S., Wells, R. and Rechsteiner, M. (1986) Science 234, 364–368.
[24] Marston, S.B. and Redwood, C.S. (1991) Biochem. J. 279, 1–16.
[25] Falk, K. and Rötzschke, O. (1993) Immunology 5, 81–94.
[26] Rammensee, H.-G., Falk, K. and Rötzschke, O. (1993) Annu. Rev. Immunol. 11, 213–244.
[27] Jardetzky, T.S., Lane, W.S., Robinson, R.A., Madden, D.R. and Wiley, D.C. (1991) Nature 353, 326–329.
[28] Hunt, D.F., Henderson, R.A.,Shabanowitz, J., Sakaguchi, K., Michel, H., Sevilir, N., Cox, A.L., Appella, E. and Engelhard, V. (1992) Science 255, 1261–1263.
[29] DiBrino, M., Parker, K.C., Shiloach, J., Turner, R.V., Tsuchida, T., Garfield, M., Biddison, W.E. and Coligan, J.E. (1994) J. Immunol. 152, 621–631.
[30] Corr, M., Boyd, L.F., Frankel, S.R., Kozlowski, S., Padlan, E.A. and Margulies, D.H. (1992) J. Exp. Med. 176, 1681–1692.