The EMBO Journal vol.11 no.8 pp.3091 - 3098, 1992

A chimeric ubiquitin conjugating enzyme that combines the cell cycle properties of CDC34 (UBC3) and the DNA repair properties of RAD6 (UBC2): implications for the structure, function and evolution of the E2s

Elizabeth T.Silver, Todd J.Gwozd, Christopher Ptak, Mark Goebl[1] and Michael J.Ellison[2]

Department of Biochemistry, University of Alberta, Edmonton, Alberta, T6G 2H7, Canada and [1]Department of Biochemistry and Molecular Biology, Indiana School of Medicine, Indianapolis, IN, USA

[2]To whom reprint requests should be sent

Communicated by J.Steitz

The CDC34 (UBC3) protein from *Saccharomyces cerevisiae* has a 125 residue tail that contains a polyacidic region flanked on either side by sequences of mixed composition. We show that although a catalytic domain is essential for CDC34 activity, a major cell cycle determinant of this enzyme is found within a 74 residue segment of the tail that does not include the polyacidic stretch or downstream sequences. Transposition of the CDC34 tail onto the catalytic domain of a functionally unrelated E2 such as RAD6 (UBC2) results in a chimeric E2 that combines RAD6 and CDC34 activities within the same polypeptide. In addition to the tail, the cell cycle function exhibited by the chimera and CDC34 is probably dependent on a conserved region of the catalytic domain that is shared by both RAD6 and CDC34. Despite this similarity, the CDC34 catalytic domain cannot substitute for the DNA repair and growth functions of the RAD6 catalytic domain, indicating that although these domains are structurally related, sufficient differences exist to maintain their functional individuality. Expression of the CDC34 catalytic domain and tail as separate polypeptides are capable of only partial function; thus, while the tail displays autonomous structural characteristics, there is considerable advantage gained when both domains coexist within the same polypeptide. The ability of these and other derivatives to restore partial function to a *cdc34* temperature-sensitive mutant but not to a disruption mutant suggests that interaction between two CDC34 polypeptides is a requirement of CDC34 activity. Based on this idea we propose a model that accounts for the initiating steps leading to multi-ubiquitin chain synthesis. Our observation that chimeric E2s can be constructed artificially that acquire new functions without loss of original functions, suggests that naturally occurring E2s are candidates for rapid evolutionary change. Finally, other evidence suggests a function for tails that has not been previously reported. Similar findings to those described here are reported in an accompanying manuscript (Kolman *et al.*, 1992).

Key words: cell cycle/CDC34 (UBC3)/RAD6 (UBC2)/ubiquitin conjugation/yeast

Introduction

The covalent attachment of ubiquitin (Ub) to either damaged or naturally short-lived proteins is the first step in a series of events leading ultimately to degradation, a process which is now widely believed to be involved in a diverse range of functions including cell division cycle regulation, gene expression, the stress response and DNA-damage processing. The transfer of Ub is mediated by a class of enzymes referred to as the Ub conjugating enzymes, or E2s (reviewed in Pickart, 1988; Jentsch *et al.*, 1990). While structural homologues of these enzymes have been identified in higher organisms, molecular biological studies of the yeast *Saccharomyces cerevisiae* have provided the greatest insight into their biological function. To date, the genes for 10 E2s have been characterized from this organism (Jentsch *et al.*, 1990; Jentsch, personal communication; Qin *et al.*, 1991). Mutational analyses of these genes indicate that the E2s perform a variety of cellular tasks. The yeast CDC34 (UBC3) enzyme which is the focus of the present work, is an essential component in the cell cycle transition from the G_1 to the S phase (Goebl *et al.*, 1988). The enzymes UBC4 and UBC5, on the other hand, play an important role in the response of cells to environmental stress (Seufert and Jentsch, 1990). In more than one case an E2 has been found to be multi-functional. The yeast RAD6 (UBC2) protein for example is required for DNA-damage processing, meiosis, sporulation and cell proliferation (reviewed in Jentsch *et al.*, 1987). Studies with temperature-sensitive mutations of RAD6 have recently shown that the cell proliferation function of RAD6 can actually be divided into a growth component and a cell cycle component, both of which are functionally unrelated to the processes cited above (Ellison *et al.*, 1991). In other recent work RAD6 has been shown to be the E2 involved in the turnover of short-lived substrates via the 'N-end rule pathway' in a manner which is apparently distinct from its other cellular roles (Dohmen *et al.*, 1991; Sung *et al.*, 1991).

Differences in E2 function must logically reflect differences in E2 structure. All E2s characterized to date have in common a conserved catalytic domain or core, of ~150 amino acid residues in length, that contains the active site cysteine residue involved in Ub transfer. Several E2s also have C-terminal extensions appended onto the catalytic domain polypeptide sequence. Unlike the catalytic domain, these extensions or tails bear little or no resemblance from one E2 to the next. The tail of the RAD6 for example is a 23 residue segment consisting mostly of polyaspartate (Reynolds *et al.*, 1985) whereas the tail of the CDC34 protein is considerably longer (125 residues) and more complex (Goebl *et al.*, 1988).

The function of these tails is far from obvious. Although it seems reasonable to expect that the tails participate either directly in substrate recognition or indirectly by interaction

with a Ub protein ligase (E3), several observations are not easily accounted for by this notion. The stress E2s UBC4 and UBC5 neither have nor require tails for normal function (Seufert and Jentsch, 1990). Furthermore, while the tail of RAD6 is needed for sporulation (Morrison *et al.*, 1988), it is not required for any of its other functions (Morrison *et al.*, 1988; Sung *et al.*, 1991). From these two examples it seems that for certain functions the catalytic domain is sufficient for normal E2 activity and that structural differences from one catalytic domain to the next are sufficient to account for differences in function. Contrary to the examples of UBC4, UBC5 and RAD6, which appear to retain function without the requirement of tails, in the present work, we have determined that the cell cycle properties of CDC34 (UBC3) are entirely dependent on its tail. We have localized the cell cycle determinant of this E2 to a 74 residue segment of the tail and have shown that this segment can function in combination with the catalytic domain of another E2 in a semi-autonomous fashion. The significance of these and other findings to the structure, function and evolution of the tails is discussed.

Results

If the tail of CDC34 were necessary for its cell cycle function then a CDC34 construct that had a tail deletion would not be expected to restore normal function to a *cdc34* mutant strain of yeast. If the tail of CDC34 were the sole functional determinant of this E2's cell cycle properties, then transposition of the tail to the catalytic domain of another E2 such as RAD6, might be able to restore CDC34 activity. We tested these ideas by creating genes for CDC34 and RAD6 that either lacked tail sequences ($CDC34\Delta_{170}$, $RAD6\Delta_{152}$) or where the tail sequences of each E2 had been transposed onto the catalytic domain of the other (CDC_{RAD}, RAD_{CDC}). Tail deletions of the *CDC34* gene were also constructed in which sequences encoding C-terminal residues up to and including the polyacidic stretch of CDC34 were eliminated ($CDC34\Delta_{244}$, $CDC34\Delta_{262}$). Diagrams of all the derivatives used in this study are shown in Figure 1. Genes were subcloned into a yeast/*Escherichia coli* expression plasmid and were tested for their ability to complement the *cdc34* temperature-sensitive (ts) mutant, YL10. At the elevated temperature, YL10 arrests cell division in the G_1 phase leading ultimately to cell death.

YL10 cells harbouring plasmids for each of the constructs described above were streaked onto duplicate plates which were then incubated at either the permissive temperature (24°C) or the non-permissive temperature (37°C). As seen in Figure 2, all E2 constructs lacking the entire CDC34 tail ($CDC34\Delta_{170}$, $RAD6\Delta_{152}$ and CDC_{RAD} and RAD6) failed to rescue cells from arrest at 37°C. Conversely, E2 constructs containing the 74 residue region of the CDC34 tail from residues 171 to 244 could fully complement the ts mutation. The possibility that the colonies arose in this and other experiments from rare recombinational events between the plasmid and the ts allele thereby resurrecting a functional gene was ruled out by the frequency of cell viability at the elevated temperature (estimated at between 30 and 100%). These results indicate that the 74 residue segment of the CDC34 tail contains information that is vital for cell cycle function and that the polyacidic stretch and downstream peptide sequence are not essential for CDC34 function. The observation that the CDC34 tail confers normal

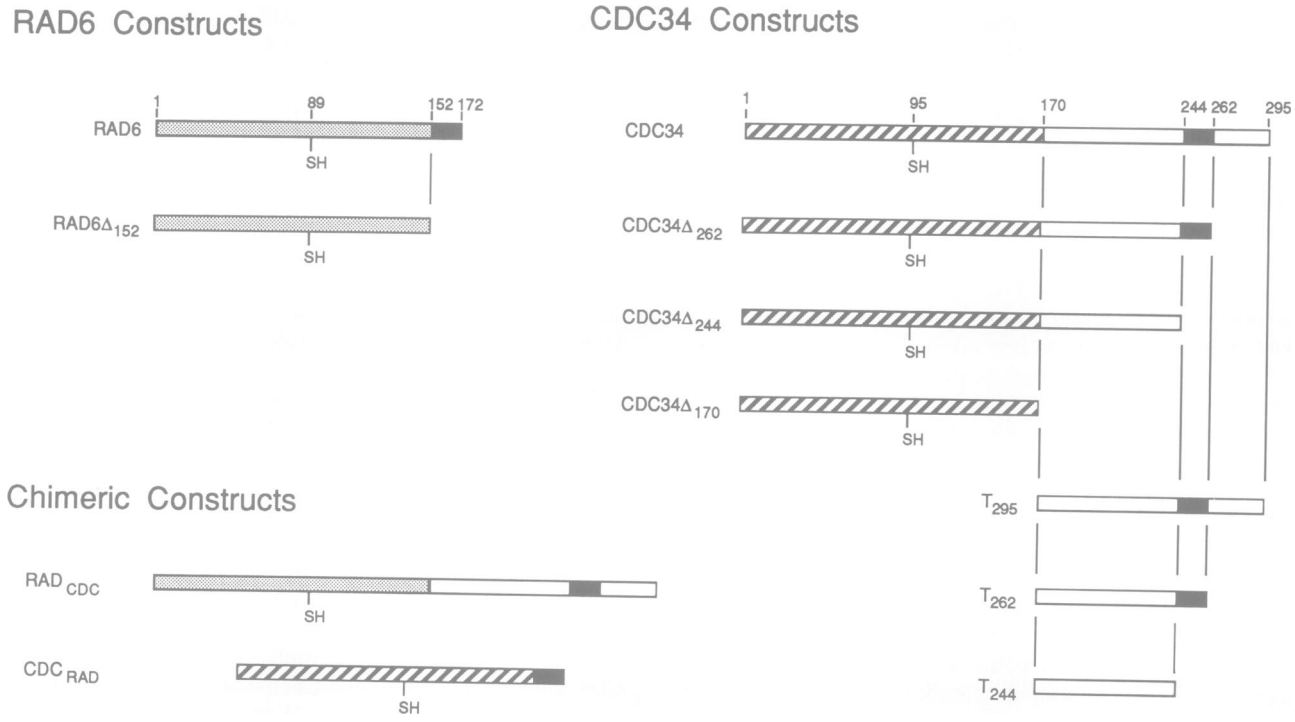

Fig. 1. Ubiquitin conjugating enzyme variations. The primary structures of CDC34 (UBC3), RAD6 (UBC2), and each of their variations used in this study are shown. Numbers refer to specific amino acid residues that either constitute important landmarks (such as the active site cysteine) or demarcate the N-terminal side of deletion boundaries and protein fusion junctions. The CDC34 and RAD6 catalytic domains are represented by striped and grey boxes respectively. The CDC34 tail is shown in white, and contains an acidic region, marked in black. The acidic tail of RAD6 is also shown in black.

cell cycle function on the RAD_{CDC} chimeric E2 illustrates that the tail and not the catalytic domain is the central determinant in defining the unique attributes of CDC34.

The growth complementation observed in Figure 2 for all constructs containing the critical region of the CDC34 tail could arise for one of two reasons. First, the CDC34 ts polypeptide is entirely non-functional, therefore complementation is entirely dependent on the E2 construct in question. Second, normal CDC34 function requires the combined activity of two active domains, one of which is supplied from a functional portion of the CDC34 ts polypeptide, whereas the other function is provided by a domain contained on the plasmid construct. To discern between these two possibilities, a cdc34 mutant strain was constructed carrying a disruption at the CDC34 locus (proximally located to the active site cysteine). This strain was rendered viable by the presence of a wild-type *CDC34* gene contained on a *URA3*-based plasmid (the maintenance plasmid). In this experiment, complementation was assessed by the ability to lose the maintenance plasmid in lieu of a functional substitute contained on a different plasmid. In the absence of selective pressure, plasmids are normally lost from cells at a frequency of between 1 in 10^3 and 1 in 10^4 cells (Koshland *et al.*, 1985). In the case of *URA3*-based plasmids, such loss events can be selected on plates containing 5-fluoroorotic acid (FOA) since failure to lose the *URA3* marker results in cell death due to FOA metabolism (Boeke *et al.*, 1987). In the case of the disruption mutant, however, loss of the *CDC34–URA3* maintenance plasmid is only permitted when its loss is compensated for by the presence of a plasmid carrying a functional CDC34 derivative and which uses a selectable marker other than *URA3*. Thus, growth of the disruption mutant on FOA plates as a consequence of the loss of the maintenance plasmid indicates that any given E2 construct described above can compensate for full loss of CDC34 function. The results of such an analysis have been combined with the results of other experiments in Table I. These results closely parallel the findings with the ts mutant (Figure 1). Only those E2 constructs containing the 74 residue CDC34 tail region were able to support growth in the presence of FOA and therefore illustrates that RAD_{CDC} and the tail-deleted derivatives $CDC34\Delta_{244}$ and $CDC34\Delta_{262}$ are capable of full complementation for loss of CDC34 mitotic function.

In view of the possibility that the cell cycle activity of the chimeric E2 RAD_{CDC}, was acquired at the expense of RAD6 function, we examined the ability of each construct to complement for the UV sensitivity phenotype of a *rad6* mutant (Figure 3). It is evident that all constructs with the RAD6 catalytic domain show strong resistance to the lethal effects of UV light whereas those constructs that use the CDC34 catalytic domain display an extreme UV sensitivity that is virtually indistinguishable from the *rad6* null mutant. These results indicate that RAD_{CDC} has acquired the mitotic cell cycle properties of CDC34 without loss of its ability to process DNA damage. Furthermore the inability of contructs that use the catalytic domain of CDC34 to complement for loss of the RAD6 UV repair function illustrates that the catalytic domains of RAD6 and CDC34 maintain aspects of functional uniqueness.

Taken together, the above experiments provide evidence that the catalytic domain and tail of CDC34 define two structural domains with a considerable measure of independence despite their physical association by a peptide bond. We examined this idea further by testing whether or not the expression of the catalytic domain and tail as separate polypeptides could restore CDC34 function in either the ts or disruption mutant. The three tail polypeptides (T_{244}, T_{262} and T_{295}) used in this experiment encompass tail residues 171–244, 171–262 and 171–295 respectively (Figure 1). As seen in Table I (rows 10–15), neither expression of these tails alone nor in combination with the CDC34 catalytic domain can complement for loss of CDC34 activity in the ts mutant at 37°C. At 34°C, however, co-expression of any one of the tail peptides in combination with the catalytic domain partially restores CDC34 function in the ts mutant. Thus, while there is unquestionable advantage to be gained by fusion of the catalytic domain and tail, the domain and 74 residue tail segment can function somewhat as discrete entities.

Interestingly, both the catalytic domains of CDC34 and RAD6 can partially complement the *cdc34* ts mutant at 34°C

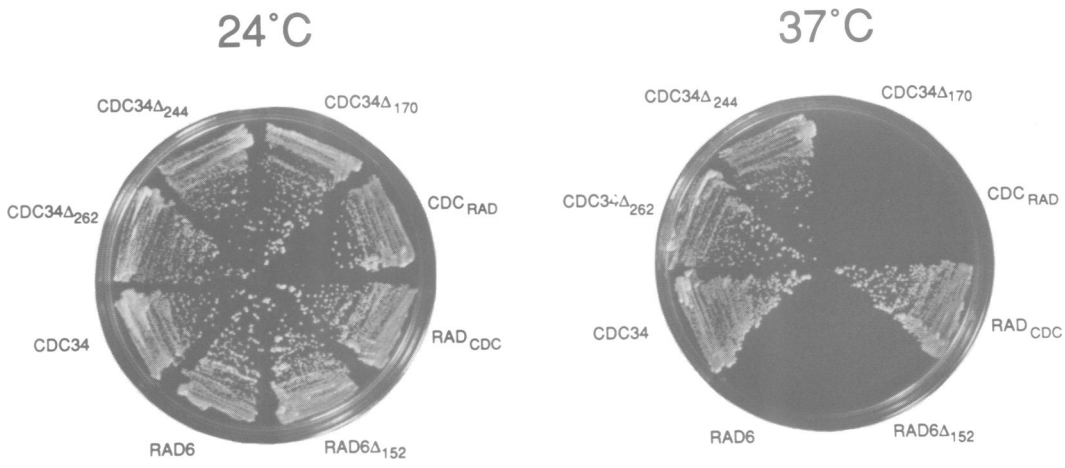

Fig. 2. The cell cycle function of CDC34 is localized to the portion of its tail between residues 171 and 244. Plasmids carrying each of the E2 constructs described in Figure 1 were transformed into the *cdc34* ts yeast mutant YL10 and tested for growth at both the permissive (24°C) and the non-permissive (37°C) temperature. Cells were streaked onto duplicate plates (see Materials and methods) and incubated at the temperatures shown.

Table I.

Plasmid	cdc34 ts mutant			cdc34 disruption mutant		
	24°C	34°C	37°C	24°C	30°C	34°C
1. Control	++++	−	−	−	−	−
2. CDC34	++++	++++	++++	++++	++++	ND
3 $CDC34\Delta_{262}$	++++	ND	++++	++++	++++	ND
4. $CDC34\Delta_{244}$	++++	ND	++++	++++	++++	ND
5. $CDC34\Delta_{170}$	++++	+	−	−	−	−
6. CDC_{RAD}	++++	ND	−	−	−	ND
7. RAD_{CDC}	++++	ND	++++	++++	++++	ND
8. RAD6	++++	−	−	−	−	−
9. $RAD6\Delta_{152}$	++++	++	−	−	−	−
10. T_{295}	++++	−	−	ND	ND	ND
11. T_{262}	++++	−	−	ND	ND	ND
12. T_{244}	++++	−	−	−	−	−
13. $CDC34\Delta_{170}$ + T_{295}	++++	++	−	ND	ND	ND
14. $CDC34\Delta_{170}$ + T_{262}	++++	++	−	ND	ND	ND
15. $CDC34\Delta_{170}$ + T_{244}	++++	++	−	−	−	−
16. RAD6 + T_{244}	++++	++	−	−	−	−
17. $RAD6\Delta_{152}$ + T_{244}	++++	++	−	−	−	−
18. UBC4	++++	−	−	ND	ND	ND
19. UBC4 + T_{244}	++++	−	−	ND	ND	ND

Functional complementation of *cdc34* mutants by E2 variants. Shown are the temperature-dependent growth characteristics of the $cdc34^{s}$ mutant YL10 or the *cdc34* disruption mutant, YES71 when carrying plasmids that encode the E2 derivatives described in Figure 1. In the case of both mutants, growth was assessed by the relative size of the colonies formed by each strain when compared with the CDC34 strain (row 2) over a suitable period of growth (3−6 days): ++++, large colonies, 1−2 mm; ++, barely discernible by eye; +, only observed with magnification; −, no growth; ND, not determined. In the case of the disruption mutant, functional complementation was assessed by the ability of strains to form colonies on FOA plates upon spontaneous loss of the *URA3−CDC34* maintenance plasmid.

in the absence of the tail constructs. In the case of $CDC34\Delta_{170}$, the degree of complementation observed is less than that of the catalytic domain and tail together. In the case of $RAD6\Delta_{152}$, however, the degree of complementation is not only greater than that observed for $CDC34\Delta_{170}$ but is not enhanced by the presence of the tail peptide T_{244}. The greater effectiveness of $RAD6\Delta_{152}$ for growth complementation relative to $CDC34\Delta_{172}$ may be due to the several-fold increase in intracellular levels of $RAD6\Delta_{152}$ compared with $CDC34\Delta_{172}$ (unpublished observations). In contrast to $RAD6\Delta_{152}$, full length RAD6 cannot complement the ts mutant indicating that the presence of the RAD6 tail has an inhibitory effect on the ability of the RAD6 catalytic domain to restore growth. Significantly, neither catalytic domain expressed alone nor in combination with T_{244} can complement the defect of the disruption mutant. Thus, all or a portion of the cdc34 ts polypeptide is also required for the restoration of function observed in each of these situations. Although the nature of the mutation conferring temperature sensitivity in the *cdc34* mutant is unknown, the data discussed above suggest it resides in the catalytic domain.

The ability of the CDC34 and RAD6 catalytic domains to restore partial function in the ts mutant is not a property shared by other E2 catalytic domains. Expression of the stress E2 UBC4 for example provides no selective advantage to the ts mutant either in the presence or absence of the tail construct T_{244} (Table I, rows 18 and 19). A sequence comparison of UBC4 with the catalytic domains of RAD6 and CDC34 (Figure 4) reveals that RAD6 and CDC34 share greater similarity to each other than to UBC4 in the C-terminal portion of the domain. It is therefore possible that this region is particularly important in defining the related cell cycle properties of RAD6 and CDC34 observed here.

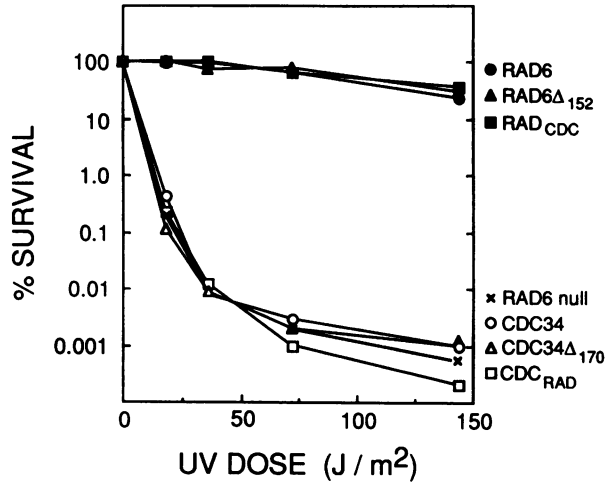

Fig. 3. CDC34 derivatives fail to complement the UV repair defect of a *rad6* mutant. A *rad6* deletion mutant carrying each of the E2 plasmids described in Figure 1 was irradiated with UV light (254 nm) as previously described (Ellison *et al.*, 1991) at the doses indicated. Cells were immediately plated and survival was determined as the percentage of surviving colonies relative to an unirradiated control for each strain.

Discussion

It is evident from the present study that both the catalytic domain and tail of CDC34 are necessary for CDC34 function, and that an advantage is gained when both are contained within the same polypeptide. While both domains are important, the cell cycle properties that make CDC34 unique from other E2s reside largely (but not exclusively) in its tail. This conclusion is based on the fact that the CDC34 catalytic domain is functionally interchangeable for the

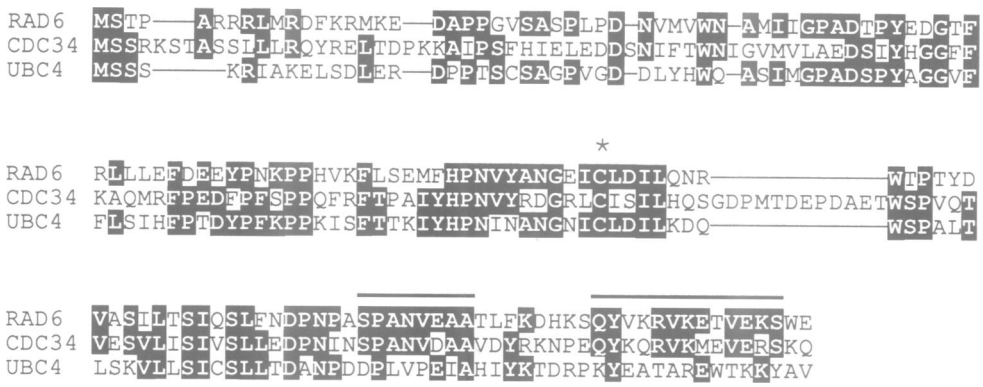

Fig. 4. E2 sequence comparison. Shown as a sequence alignment of the catalytic domains for RAD6 [UBC2; Reynolds *et al.* (1985)], CDC34 [UBC3; Goebl *et al.* (1988)] and UBC4 (Seufert and Jentsch, 1990). Identical residues are highlighted in black. The active site cysteine is marked with an asterisk. The region of similarity between RAD6 and CDC34 which is not shared by UBC4 is underlined. Alignment was obtained using the Genetic Computer Group Sequence Analysis Software package (Devereux *et al.*, 1984) on the Canadian National Research Council Molecular Biology Data Base System.

RAD6 catalytic domain, an E2 of unrelated function. A similar conclusion has been reached by Kolman *et al.* (1992).

The biochemical role of the tail is unknown. Jentsch *et al.* (1990) have speculated that tails play a role in target protein recognition, either through direct interaction or indirectly through interaction with a Ub protein ligase or E3. Another possibility, however, is discussed below. Deletional analyses presented here demonstrate that only the first 74 residues of the CDC34 tail are essential and that the remaining 51 residues (including the polyacidic region) are not required for its mitotic cell cycle function. This observation leaves open the possible role of the 51 residue segment in other less essential cellular events such as mating, meiosis, sporulation or germination.

In spite of the obvious importance of the CDC34 tail, it is also apparent that information contained within the catalytic domain itself contributes to its cell cycle activity. This conclusion is founded on the observation that the catalytic domains of RAD6 and CDC34 can both partially restore cell cycle function to the *cdc34* ts mutant whereas that of UBC4 cannot. The catalytic domains of RAD6 and CDC34 therefore share a structural relatedness that is not found in UBC4. A comparison of the amino acid sequences of these three polypeptides (Figure 4) reveals that RAD6 and CDC34 display the greatest similarity in the C-terminal portion of the catalytic domain while maintaining significant divergence from the sequence of UBC4. We therefore tentatively assign this region as another essential cell cycle determinant in addition to the tail. That RAD6 should also contain this element may prove not to be so surprising in view of its own involvement in the cell cycle at the S/G_2 phase boundary (Ellison *et al.*, 1991).

Mechanistic aspects of CDC34 function

The fact that the catalytic domains of CDC34 and RAD6 can partially restore CDC34 function in the ts mutant but not in the knockout mutant indicates that these catalytic domains must somehow utilize the tail from the ts polypeptide itself. Based on the growth advantage that is gained when the domain and tail are contained within the same polypeptide, it is difficult to imagine how the catalytic domain and ts polypeptide could co-operate with one another unless each was held in close proximity to the other either by their direct interaction or by a third protein such as an E3. One conceivable spatial arrangement between the plasmid-encoded catalytic domain and the tail of the ts polypeptide is that the plasmid catalytic domain physically displaces a dysfunctional catalytic domain contained on the ts polypeptide and assumes a position with respect to the ts polypeptide tail that sufficiently mimics the structure of normal CDC34 so as to partially restore CDC34 function. Sterically, such an interaction would appear unlikely. Furthermore, were this interaction to occur, then expression of the catalytic domain and tail as separate polypeptides would be expected to complement for CDC34 function in the disruption mutant, which it does not (Table I, row 15).

An alternative spatial arrangement that we favour is that full CDC34 function is normally dependent upon the close proximity of two CDC34 monomers. If this proximity were to be based upon the direct interaction of one monomer with the other, then the catalytic domains of RAD6 and CDC34 could function to partially stabilize the temperature-sensitive region of the ts polypeptide through direct interaction. In addition to the CDC34 tail, the C-terminal region of the RAD6 and CDC34 catalytic domains may also participate in such an interaction in view of the high degree of similarity that RAD6 and CDC34 share in this region. Chromatographic evidence obtained in V.Chau's laboratory (personal communication) and our own (unpublished) indicates that CDC34 associates to form a dimer *in vitro*. These observations coupled with the *in vivo* observations reported here would argue that dimer formation is an important aspect of CDC34 function *in vivo*.

The assembly of a multi-Ub chain onto a substrate that has been targeted for degradation appears to be a prerequisite for its turnover (Hershko and Heller, 1985; Bachmair and Varshavsky, 1989; Chau *et al.*, 1989; Gregori *et al.*, 1990). CDC34 and $E2_{25K}$ (from rabbit reticulocytes) are both capable of multi-Ub chain assembly *in vitro* in an E3 independent manner (Chen and Pickart, 1990; Haas *et al.*, 1991). In each case, the link that couples one Ub to the next is an isopeptide bond formed at Lys48. The interaction of one E2 molecule with another, either of their own accord, or through a third protein would also naturally position one Ub molecule proximal to another and might therefore explain how this link is formed. Figure 5 shows a model of an E2 dimer in which each monomer is coupled at its active site cysteine with Ub. The dimer is represented with two-fold

rotational symmetry which positions the C terminus of each Ub molecule close to Lys48 of its neighbour. In this model, the substrate has been arbitrarily positioned. While the model in Figure 5 establishes a plausible geometry of components for linking Ub to the substrate and initiating chain assembly, obviously nothing can be said at the present time about the sequence of events that brings this about, or the mechanism associated with chain extension beyond this point.

The observation that the RAD6 core functions in combination with the CDC34 ts polypeptide raises the intriguing possibility that dimer geometry is not restricted to the symmetrical homo variety as illustrated in Figure 4 but could consist of mixed complexes composed of different E2s.

Evolution of the Ub conjugating enzymes

A comparison of the sequences of various E2s indicates that these enzymes have most probably evolved by two distinct routes. On the one hand, the catalytic domain has evolved slowly, largely through amino acid replacement. This is not surprising in view of the constraints placed upon its structure by the multiple protein interactions that all such enzymes hold in common. These include interaction with Ub, interaction with Ub activating enzyme (E1) and interactions with E3. On the other hand, the observations that tails are not a feature common to all E2s and that no two tails look much alike, argues that the tails have evolved more recently and more abruptly than the catalytic domain. Human (Schneider *et al.*, 1990; Koken *et al.*, 1991a), *Drosophila* (Koken *et al.*, 1991b) and *Schizosaccharomyces pombe* (Reynolds *et al.*, 1990) homologues of RAD6 for example share considerable similarity with the catalytic domain of their yeast counterpart (upwards of 70% identity) yet neither has the polyacidic tail found on RAD6. Conversely, although RAD6 and the wheat germ Ub conjugating enzyme, $E2_{22K}$ (Sullivan and Vierstra, 1989) are only distally related, $E2_{22K}$ also has a polyacidic tail which almost certainly was not acquired from any common progenitor to RAD6 or $E2_{22K}$. Therefore, it is possible that two catalytic domains, each specifying its own unique function, may share common functions by the acquisition of similar tails by different means.

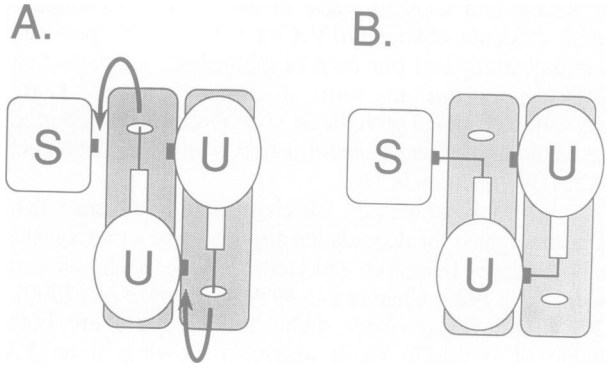

Fig. 5. A model for multi-ubiquitin chain initiation. **A.** shows a complex formed between the substrate to be ubiquitinated (S) and an E2 dimer composed of two E2 monomers (shaded boxes) each coupled at its active site (white ellipse) to Ub (U). An acceptor lysine on the substrate and Lys48 of Ub have been represented by black boxes. **B.** shows the initiation of chain synthesis as a consequence of isopeptide bond formation.

One way in which certain E2s may have evolved tails is by a fortuitous recombinational event that fused the coding sequence of the E2 in question to a sequence that encoded a domain of an unrelated polypeptide. Under some circumstances, a given domain may normally have interacted with a third protein. In turn, the newly gained proximity between this protein and the catalytic domain of the E2 may have acted as the catalyst for its ubiquitination. If even a slight selective advantage were forthcoming from such an event then an initially inefficient reaction could be optimized by the gradual co-evolution of E2 and substrate alike. In such a scheme, the role of an E3 is unclear since target specificity is naturally expected to reside in the tail of the E2. Another possibility is that a newly appended domain contained a site of interaction with itself thereby facilitating dimer formation and as a consequence setting into motion the evolutionary events that resulted in multi-Ub chain formation as discussed above.

In other cases, tails may have evolved by a mutation close to the termination codon of a particular E2 that extended the open reading frame into the untranslated portion of the gene. Since in this instance the tail would not have been derived from a pre-existing structure, the possible selective advantage of this event is less easily rationalized. One way in which the acquisition of this type of tail might provide an instant selective advantage is highlighted by the results shown in Table I. RAD6 without its tail is able to partially complement for a CDC34 deficiency whereas RAD6 with its tail cannot. The most reasonable explanation for this result is that the tail of RAD6 prevents the interaction of the catalytic domain with some component required in the CDC34-related function. In this particular case the absence of the RAD6 tail is beneficial. In general, however, the prospect that an E2 with one particular function can cross interact with components that are required for another E2s function, could be detrimental. Dohman *et al.* (1991) have observed for instance that the turnover rate of a short-lived test protein actually increases by 4-fold in a *rad6* mutant background, a result which suggests that the targeting of a particular substrate by one E2 can be attenuated by competition with another. Some tails therefore, may fulfil an inhibitory role, providing a way of functionally isolating selected E2s from interfering with the behaviour of other E2s. The idea of the RAD6 tail as an attenuator of protein−protein interaction is diametrically opposed to the notion that the tail is involved in direct substrate recognition as proposed by others (Sung *et al.*, 1988).

Whatever the mechanism by which tails were acquired, or whatever the selective pressures that may have assured their survival, the acquisition of tails undoubtedly provided E2's with the means to evolve new functions over a relatively brief period of time. Central to this idea, however, is the constraint that old functions are not sacrificed for the acquisition of new functions; or in other words that the appendage of a tail onto an E2 did not interfere with the E2's original function. That this condition can be met is aptly illustrated by the ability of RAD6 to acquire the full mitotic cell cycle properties of CDC34 upon receipt of the CDC34 tail without loss of its own DNA repair functions. The potential plasticity that is imparted on these enzymes by the evolutionary mechanisms discussed above, suggests that the spectrum of E2 function observed in simpler eukaryotes represents only

a fraction of the possible functions that will be found in more complex organisms.

Materials and methods

Yeast expression plasmid construction

Here, sufficient information is provided for the unambiguous sequence assignment of each expression plasmid used in the present work. Further details on the construction of these plasmids is available upon request. The DNA sequence of each E2 derivative is ultimately derived from the previously determined sequences for either RAD6 (Reynolds *et al.*, 1985), CDC34 (Goebl *et al.*, 1988) or UBC4 (Seufert and Jentsch, 1990) in addition to the modification described below.

Gene cassettes encoding the various E2 derivatives shown in Figure 1 were constructed from appropriate templates using the polymerase chain reaction (PCR) in combination with primers that defined the coding sequence boundaries of each polypeptide. Each cassette was designed to precisely replace the Ub gene cassette found in the yeast expression plasmid YEp96 (Ellison and Hochstrasser, 1991) such that the start and stop codons of each gene occupy identical positions with respect to the invariant non-coding flanking sequences that constitute the *CUP1* promoter and *CYC1* transcriptional terminator of YEp96. YEp96 is a derivative of YEp46 (Ecker *et al.*, 1987), a high copy yeast/*E.coli* shuttle plasmid carrying the *TRP1* auxotrophic selectable marker.

To facilitate further manipulation, selected restriction enzyme sites were introduced into the DNA coding sequence of each derivative that in some cases resulted in limited changes to the peptide sequence. The second and third codon of each derivative were changed to create a *Sac*I site. The consequence of this alteration is that each polypeptide initiates with the sequence Met-Ser-Ser. This gene modification has no effect on the peptide sequence of wild-type CDC34 or its derivatives CDC34Δ_{170}, CDC34Δ_{244}, CDC34Δ_{262} and CDC$_{RAD}$, but does result in the conservative replacement of Thr for Ser at position 3 of RAD6, RAD6Δ_{152} and RAD$_{CDC}$. This replacement has no effect on normal RAD6 function (Ellison *et al.*, 1991). For the CDC34 tail derivatives T_{244}, T_{262} and T_{295}, the presence of the *Sac*I site appends the sequence Met-Ser-Ser to Ile-171 (the designated start of the CDC34 tail domain).

In the case of *RAD6* and *CDC$_{RAD}$*, and *Eco*RV site positioned at the boundary between the catalytic domain and tail coding sequences results in a Met-153 to Ile substitution at the beginning of the RAD6 tail. Based on unpublished data and data presented here, this replacement has no effect on RAD6 function.

A control plasmid (pES13) carrying no E2 sequence was constructed by deleting the Ub coding sequence of YEp96 between the internal *Bgl*II and *Sal*I sites. *LEU2* counterparts of the RAD6Δ_{152}, RAD6, CDC34Δ_{170} and UBC4 expression plasmids were constructed by replacing the *TRP1* marker situated between the *Hpa*I and *Cla*I sites, with the *Hpa*I–*Nar*I *LEU2* fragment from YEp351 (Hill *et al.*, 1986).

The *URA3* counterpart of the *TRP1-CDC34* expression plasmid was constructed by replacing the *TRP1* marker situated between the *Hpa*I and *Cla*I sites with the *Hpa*I/*Nar*I *URA3* fragment from YEp352 (Hill *et al.*, 1986).

The DNA sequences of all plasmid-borne gene cassettes described above were verified using the double-stranded DNA chain termination method either manually or by an Applied Biosystems automated DNA sequencer (model 373A) operated by the Department of Biochemistry Facility for DNA synthesis and sequencing at the University of Alberta.

Yeast strains

All *TRP1* based plasmids described above were transformed either singly or in combination with the *LEU2* based plasmids into the yeast strain YL10 (genotype: *MATα, ura3-52, trp1Δ63, leu2Δ1, his3Δ, cdc34-2*). YL10 carries a temperature-sensitive mutation in the *CDC34* allele.

A mutant of YL10 (YES71) in which the *cdc34-2* locus was disrupted by *HIS3* (interrupting codons 104 and 105 which are proximal to the active site codon, Cys-95) was constructed using a previously described strategy (Goebl *et al.*, 1988). The growth of this mutant is maintained by the presence of the *URA3-CDC34* expression vector described above.

For plasmid loss experiments, YES71 was singly transformed with either the *LEU1* expression plasmids for *RAD6, RAD6Δ$_{152}$, CDC34, CDC34Δ$_{170}$* and *UBC4* or the *TRP1* plasmid for T_{244}, or the *TRP1* control plasmid (pES12). YES71 was also co-transformed with the *LEU2* plasmid versions of *RAD6, RAD6Δ, CDC34Δ$_{170}$* and *UBC4* in combination with the *TRP1-T_{244}* plasmid.

For UV lethality experiments, the *RAD6* deleted strain, KMY20 (genotype: *MATα, ade2-101, his3-832, trp1-289, ura3-52 rad6Δ::URA3*, a gift from K.Madura, MIT) were individually transformed with the *TRP1* based plasmids carrying *RAD6Δ$_{152}$, RAD6, CDC34, CDC34Δ$_{170}$, RAD$_{CDC}$* and *CDC$_{RAD}$*.

Media and growth conditions

Yeast were grown in liquid culture or on 1.5% agar plates using defined minimal media (Sherman *et al.*, 1981) supplemented with appropriate combinations of uracil, adenine, histidine, leucine and tryptophan at concentrations previously specified (Sherman *et al.*, 1981) as required for plasmid selection and according to strain genotype.

For plasmid loss experiments, 1×10^6 cells were plated onto minimal media agar plates containing 1 g/l FOA (Sigma), uracil, histidine, leucine and tryptophan. Under these conditions, loss of the *URA3-CDC34* maintenance plasmid giving rise to colonies occurred at a frequency of 1/1000–1/5000 cells. In all experiments, cells were plated from exponentially growing cultures at 24°C in liquid medium.

Acknowledgements

We thank K.S.Ellison for editorial comments and S.Smith for secretarial assistance. This work was made possible by grants from the Medical Research Council of Canada, the Alberta Heritage Foundation for Medical Research and the National Cancer Institute of Canada. M.J.E. is an MRC Research Scholar.

References

Bachmair,A. and Varshavsky,A. (1989) *Cell*, **56**, 1019–1032.
Boeke,J.D., Trueheart,J., Natsoulis,G. and Fink,G.R. (1987) *Methods Enzymol.*, **154**, 164–175.
Chau,V., Tobias,J.W., Bachmair,A., Marriott,D., Ecker,D.J., Gonda,D.K. and Varshavsky,A. (1989) *Science*, **243**, 1576–1583.
Chen,Z. and Pickart,C.M. (1990) *J. Biol. Chem.*, **265**, 21835–21842.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acid Res.*, **12**, 387–395.
Dohmen,R.J., Madura,K., Bartel,B. and Varshavsky,A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 7351–7355.
Ecker,D.J., Khan,M.I., Marsh,J., Butt,T.R. and Crooke,S.T. (1987) *J. Biol. Chem.*, **262**, 3524–3527.
Ellison,M.J. and Hochstrasser,M. (1991) *J. Biol. Chem.*, **266**, 21150–21157.
Ellison,K.S., Gwozd,T., Prendergast,J.A., Paterson,M.C. and Ellison,M.J. (1991) *J. Biol. Chem.*, **266**, 24116–24120.
Goebl,M.G., Yochem,J., Jentsch,S., McGrath,J.P., Varshavsky,A. and Byers,B. (1988) *Science*, **241**, 1331–1335.
Gregori,L., Poosch,M.S., Cousins,G. and Chau,V. (1990) *J. Biol. Chem.*, **265**, 8354–8357.
Haas,A., Reback,P.M. and Chau,V. (1991) *J. Biol. Chem.*, **266**, 5104–5112.
Hershko,A. and Heller,H. (1985) *Biochem. Biophys. Res. Commun.*, **128**, 1079–1086.
Hill,J.E., Myers,A.M., Koerner,T.J. and Tzagoloff,A. (1986) *Yeast*, **2**, 163–167.
Jentsch,S., McGrath,J.P. and Varshavsky,A. (1987) *Nature*, **329**, 131–134.
Jentsch,S., Seufert,W., Sommer,T. and Reins,H.-A. (1990) *Trends Biol. Sci.*, **15**, 105–108.
Koken,M.H.M., Reynolds,P., Jaspers-Dekker,I., Prakash,L., Prakash,S., Bootsma,D. and Hoeijmakers,J.H.J. (1991a) *Proc. Natl. Acad. Sci. USA*, **88**, 8865–8869.
Koken,M.H.M., Reynolds,P., Bootsma,D., Hoeijmakers,J.H.J., Prakash,S. and Prakash,L. (1991b) *Proc. Natl. Acad. Sci. USA*, **88**, 3832–3836.
Kolman,C.J., Toth,J. and Gonda,D.K. (1992) *EMBO J.*, **11**, 3081–3090.
Koshland,D., Kent,J.C. and Hartwell,L.H. (1985) *Cell*, **40**, 393–403.
Morrison,A., Miller,E.J. and Prakash,L. (1988) *Mol. Cell. Biol.*, **8**, 1179–1185.
Pickart,C.M. (1988) In Rechsteiner,M. (ed.), *Ubiquitin*. Plenum Publishing Co., New York, pp. 77–99.
Qin,S., Nakajima,B., Nomura,M. and Arfin,S. (1991) *J. Biol. Chem.*, **266**, 15549–15554.
Reynolds,P., Weber,S. and Prakash,L. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 168–172.

Reynolds,P., Koken,M.H.M., Hoeijmakers,J.H.J., Prakash,S. and Prakash,L. (1990) *EMBO J.*, **9**, 1423–1430.
Schneider,R., Eckerskorn,C., Lottspeich,F. and Schweiger,M. (1990) *EMBO J.*, **9**, 1431–1435.
Seufert,W. and Jentsch,S. (1990) *EMBO J.*, **9**, 543–550.
Sherman,F., Fink,G.R. and Hicks,J.B. (1981) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sullivan,M.L. and Vierstra,R.D. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9861–9865.
Sung,P., Prakash,S. and Prakash,L. (1988) *Genes Dev.*, **2**, 1476–1485.
Sung,P., Berleth,E., Pickart,C., Prakash,S. and Prakash,L. (1991) *EMBO J.*, **10**, 2187–2193.

Received on March 26, 1992; revised on May 5, 1992