The HCF repeat is an unusual proteolytic cleavage signal

Angus C. Wilson,1 Michael Gregory Peterson,2 and Winship Herr1

1Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724 USA; 2Tularik Inc., South San Francisco, California 94080 USA

The herpes simplex virus VP16-associated protein HCF is a nuclear host-cell factor that exists as a family of polypeptides encoded by a single gene. The mature HCF polypeptides are amino- and carboxy-terminal fragments of a large ~300-kD precursor protein that arise through cleavage at one or more centrally located sites. The sites of cleavage are the HCF repeats, highly conserved 26-amino-acid sequences repeated six times in the HCF precursor protein. The HCF repeat alone is sufficient to induce cleavage of a heterologous protein, and cleavage occurs at a defined site—PPCE/THET—within the HCF repeat. Alanine-scan mutagenesis was used to identify a large 18-amino-acid segment of the HCF repeat that is important to induce cleavage of a heterologous protein. Even though HCF is cleaved, the majority of amino- and carboxy-terminal cleavage products remain tightly, albeit noncovalently, associated. Modulation of this noncovalent association may provide a mechanism for regulating HCF activity. For example, the cleaved products of an alternative mRNA splicing variant of HCF do not remain associated.

[Key Words: Multiprotein–DNA complex; transcription; VP16; herpes simplex virus]

Received August 2, 1995; revised version accepted August 30, 1995.

In eukaryotes, the activity of intracellular proteins is often modulated through post-translational modification. In some cases, the modification is reversible, and the activity of a protein can cycle between different states, as with phosphorylation [see Hunter 1995]. In contrast, with proteolysis, the post-translational modification is essentially irreversible. Sequence-specific cleavage to regulate the activity of extracellular proteins is common. A well-characterized example is the serine proteases that form the blood clotting cascade [Jackson and Nemerson 1980]. These proteases recognize and precisely cleave at sequences of ~4–6 residues within their target proteins. In contrast, less is known about intracellular proteins, particularly nuclear proteins, that are cleaved at specific sites.

Here, we describe an unusual form of nuclear proteolysis in which the substrate, the human HCF protein, is cleaved at a series of relatively large 26-amino-acid repeats in the protein. HCF [for herpes simplex virus host cell factor, but also referred to as C1 [Kristie et al. 1989], VCAF [Xiao and Capone 1990], and CFF [Katan et al. 1990]], is a cellular activity required for transcriptional activation of herpes simplex virus [HSV] immediate early [IE] promoters by the viral transcriptional activator protein VP16 [Gerster and Roeder 1988; Kristie et al. 1989; for review, see Thompson and McKnight 1992; O’Hare 1993]. VP16 [also known as Vmw65 or α-TIF] is a HSV virion protein that is released into the cell upon infection and forms a heterodimeric protein complex with HCF. The association of VP16 with HCF promotes formation of a multiprotein complex with the POU domain-containing transcription factor Oct-1 on the cis-regulatory target of VP16 activation, the TAATGARAT motif [R designates a purine] found in HSV IE promoters. This VP16-induced complex can then initiate HSV transcription by virtue of a very potent transcriptional activation domain in VP16.

The cellular function of HCF is unknown but is probably important because HCF activity has been well conserved during evolution. In addition to mammalian cells, nematode and insect cells contain a HCF activity able to promote VP16-induced complex formation [Kristie et al. 1989; Wilson et al. 1993a]. One hypothesis is that the requirement for VP16 association with HCF before activation of HSV IE gene expression serves as a mechanism to gauge the proliferative status of the infected cell. Proliferative potential may influence whether the viral infection follows a lytic or latent pathway. Consistent with this hypothesis, HCF expression is highest in embryonic tissues, which have many proliferating cells, and is low in most adult tissues examined, which have many resting cells [Wilson et al. 1995].

The human HCF gene directs the synthesis of a large transcript coding for a 2035-amino-acid protein with an apparent molecular mass of ~300 kD [Wilson et al. 1993b; see also Kristie et al. 1995]. The most striking feature of the sequence of this protein is a set of six near-perfect 26-amino-acid repeats, referred to as HCF repeats [Wilson et al. 1993b], located near the middle of the HCF open reading frame [ORF]. Despite being en-
Wilson et al.

coded by a large ORF, HCF isolated from HeLa cells consists of a family of many different polypeptides ranging in size from 110 to 300 kD [Kristie and Sharp 1993; Wilson et al. 1993b]. Our previous study suggested that the smaller HCF polypeptides arise through proteolysis of the large 300-kD protein, called HCF\textsubscript{300} [also referred to as a 235-kD protein [Kristie et al. 1995]], but the mechanism or signals responsible for the processing of HCF were not revealed [Wilson et al. 1993b].

Here we show that HCF is processed via proteolytic cleavage at the HCF repeats. The signal for proteolysis is a large 18-amino-acid segment of the highly conserved HCF repeat sequence. Following cleavage, the majority, but not all, of the amino- and carboxy-terminal HCF fragments remain stably but noncovalently associated with one another.

**Results**

Figure 1 shows a diagram of the recombinant human HCF (rHCF) used in our experiments. HCF consists of regions enriched in particular amino acids [e.g., basic and acidic residues], but the most prominent feature is the centrally located HCF repeats of which there are six near-perfect copies (HCF repeats 1–3, 5, 6, and 8; solid arrowheads) and two imperfect copies (HCF repeats 4 and 7; open arrowheads). For the experiments described here, we modified the HCF protein by including either a flu hemagglutinin (HA) or a bacteriophage T7 gene 10 (T7) epitope tag at the amino terminus. These modifications allowed us to distinguish transiently expressed recombinant HCF from endogenous HCF.

Metabolic labeling reveals the cellular location, time course, and multiplicity of HCF processing

To study the processing of HCF in vivo, we performed a pulse-chase experiment. Over a 20-hr time course, we assayed the appearance and disappearance of recombinant HCF polypeptides that were metabolically labeled with [35\textsuperscript{S}]methionine during a 30-min pulse labeling of 293 cells. To characterize the subcellular localization of precursor and mature forms of HCF, cell extracts were divided into crude nuclear and cytoplasmic fractions.

The protein extracts were subsequently denatured, and amino-terminal rHCF polypeptides selectively recovered by immunoprecipitation and resolved by SDS–polyacrylamide gel electrophoresis.

Figure 2 shows the results of such an experiment. We identify five HCF-related bands: The band resulting from the slowest mobility polypeptide corresponds to the HCF\textsubscript{300} polypeptide; the four additional bands (labeled A–D) represent processed HCF fragments. Results described below indicate that each of bands B and C contain two processed HCF fragments of similar size. Figure 2B shows quantitation of the levels of HCF\textsubscript{300} in the nuclear and cytoplasmic fractions, and the level of the processed HCF fragment A in the nuclear fraction.

Immediately after labeling for 30 min, HCF\textsubscript{300} was the predominant labeled polypeptide in both the cytoplasmic [lane 2] and nuclear [lane 8] fractions, suggesting that HCF\textsubscript{300} is the primary HCF translation product. In the cytoplasmic fractions, HCF\textsubscript{300} became progressively less abundant during the chase period, during which the labeled cells were incubated with unlabeled methionine for 1–20 hr [Fig. 2A, lanes 3–6; see Fig. 2B]. In contrast, the level of HCF\textsubscript{300} polypeptide in the nuclear fraction increased slightly during the first hour of the chase [Fig. 2A, cf. lanes 8 and 9; see Fig. 2B], and then subsequently diminished in parallel with HCF\textsubscript{300} in the cytoplasmic fraction [lanes 9–12]. The early increase in HCF\textsubscript{300} accumulation in the nuclear fraction may reflect continued transport of recently synthesized HCF precursor from the cytoplasm to the nucleus during this time period.

In the nuclear fractions, the progressive loss of HCF\textsubscript{300} was paralleled by an accumulation of the smaller amino-terminal HCF fragments particularly fragment A and one or both of the fragments in band B [Fig. 2A, see lanes 11, 12 and A; Fig. 2B, Frag. A]. In this experiment, the half-life of nuclear HCF\textsubscript{300} was ~3 hr, indicating a relatively slow time course of HCF processing. The inverse relationship between the loss of HCF\textsubscript{300} and accumulation of smaller HCF fragments indicates that the smaller HCF polypeptides arise by proteolytic processing of HCF\textsubscript{300} in the nucleus. Over time, the intermediate-sized HCF fragments in bands C and D in the nuclear fraction appear and disappear as the smallest HCF fragments in bands A and B continue to accumulate [cf. lanes 3–6].

![Figure 1. Schematic representation of the full-length epitope-tagged rHCF ORF. The 2035 codon HCF ORF is represented by the rectangular box. The six near-perfect HCF repeats (solid arrowheads) and two degenerate HCF repeats (open arrowheads) are indicated. Regions generally enriched in particular types of amino acids are indicated by solid and shaded boxes. WYF signifies tryptophan, tyrosine, and phenylalanine. The position of the alternative splice that removes codons 382–450 is shown. Epitope tags recognized by either αHA (12CA5) or αT7.tag (Novagen Inc.) monoclonal antibodies are located at the amino terminus (αHCF\textsubscript{N} antibodies). Monoclonal antibody αHCF-M2 [Wilson et al. 1995] recognizes an epitope specified by codons 1436–1508, whereas rabbit polyclonal αHCF-H12 antiserum was prepared against amino acids 1508–2024 (αHCF, Wilson et al. 1993b).](image-url)
The HCF repeat signals proteolytic cleavage

10–12]. Apparently, the intermediate-sized fragments continue to be proteolytically processed to generate the smallest fragments, suggesting that HCF can undergo multiple rounds of processing within the nucleus.

In addition to proteolytic cleavage, the experiment in Figure 2A suggests an additional post-translational modification of HCF, because the electrophoretic mobility of HCF300 and the cleaved HCF fragments decreases progressively over time (Fig. 2A, cf. lanes 2–6 and 8–12). We do not know the reason for the decreased mobility, but it may be attributable to glycosylation [Jackson and Tjian 1988] or phosphorylation. Consistent with the former possibility, HCF is glycosylated [Wilson et al. 1993b] and O-linked N-acetylglucosamine residues can be labeled with radiolabeled galactose [A.C. Wilson and K. LaMarco, unpubl.].

Mutations in HCF repeats 1 and 2 disrupt formation of the smallest amino-terminal HCF fragments

The number and apparent size of the processed HCF fragments suggest that the HCF precursor is cleaved at multiple sites toward the center of the polypeptide. A good candidate for a cleavage signal is the HCF repeat, which is present in multiple copies near the center of the ORF (Fig. 1). To test this possibility, we assayed the effects on HCF processing of either deleting the entire HCF repeat region or mutating the two amino-terminal HCF repeats 1 and 2, individually or in combination, as shown in Figure 3. Figure 3B shows the sequence of HCF repeats 1 and 2 and the five alanine substitutions (called KO for knockout) used to disrupt their function. To assay HCF processing, epitope-tagged full-length and mutated forms of rHCF were expressed transiently in 293 cells. Amino-terminal rHCF polypeptides were subsequently recovered by immunoprecipitation and detected by immunoblotting with the same epitope tag-specific antibody.

Figure 3C shows the result of such an experiment. As expected, full-length HCF [rHCFFL] generated at least six amino-terminal epitope-tagged polypeptides (indicated with dots; lane 2). In contrast, expression of the deleted form of HCF [rHCFFLΔrep; lane 6] resulted in a single major new epitope-tagged polypeptide of ~200 kD. [One minor ~120-kD species is not observed consistently and migrates slightly faster than the smallest
Figure 3. Proteolytic cleavage requires the HCF repeats. (A) The rHCF ORF and derivatives are shown schematically. In rHCFΔrep, residues 1012–1435 are deleted, removing all of the HCF repeats. In rHCFrepl(KO) and rHCFrep2(KO), five conserved residues of repeats 1 and 2, respectively, have been substituted with alanine. rHCFrepl+2(KO) combines both substitution mutants. Each polypeptide includes an amino-terminal α7-epitope tag. (B) Sequences of HCF repeats 1 and 2 showing the 5 residues that were substituted with alanines to generate the repl(KO) and rep2(KO) derivatives. (C) Extracts were prepared from mock-transfected 293 cells (lane 1) or cells transfected with 5 μg of expression plasmids encoding (rα7-epitope-tagged rHCF, (lane 2), rHCFrepl(KO) (lane 3), rHCFrep2(KO) (lane 4), rHCFrep1+2(KO) (lane 5), and rHCFΔrep (lane 6). Recombinant HCF proteins were recovered by immunoprecipitation using the α7-tag antibody, resolved on a 7% SDS-polyacrylamide gel, and detected by immunoblotting with the α7-tag antibody. The dots indicate specific epitope-tagged products generated by the wild-type rHCF expression construct. The mobility of prestained molecular mass markers is indicated in kD.

rHCFΔrep fragment. This species probably arises from non-specific HCF proteolysis. The smaller size of the ~200-kD rHCFΔrep polypeptide, compared to HCF300, is appropriate for the primary translation product from this HCF deletion construct. Significantly, the shorter products characteristic of HCF processing were not detected (Fig. 3C, cf. lanes 2 and 6), indicating that the region of HCF containing the HCF repeats is necessary for proper processing.

To determine whether individual HCF repeats are involved in HCF processing, we analyzed processing of the HCF KO substitution mutants rHCFrepl(KO), rep2(KO), and rep1+2(KO). As shown in Figure 3C (lanes 3–5), expression of these mutant HCF polypeptides generated different patterns of processed fragments compared to the pattern generated by wild-type rHCFFL protein (lane 2). Mutation of HCF repeat 1 (lane 3), which is located closest to the amino terminus of HCF, resulted in the absence of the smallest epitope-tagged polypeptide, whereas mutation of HCF repeat 2 (lane 4) did not disrupt formation of the smallest HCF fragment but, instead, resulted in loss of the second smallest HCF polypeptide (cf. lanes 2 and 4). Combining the mutations in HCF repeats 1 and 2 [rHCFFLrep1+2(KO)] resulted in loss of both of the smaller HCF polypeptides (lane 5).

These results indicate that the HCF repeats induce HCF cleavage. The number and size of amino-terminal fragments also suggest that each of the near-perfect HCF repeats directs cleavage at or near the repeat itself. For example, the relative sizes of the three smallest HCF fragments (Fig. 3C, lane 2) mirror the larger and closer spacing between repeats 1 and 2, and 2 and 3, respectively (see Fig. 3A).

A single HCF repeat is sufficient to induce cleavage of a heterologous protein

Having demonstrated that the HCF repeats are required for proper processing of HCF, we next determined whether the HCF repeat itself is sufficient to signal proteolytic processing. For this purpose, we assayed the ef-
fects of placing a single HCF repeat into human Oct-1, a very different but well-characterized nuclear protein. A single 35-codon fragment encoding a single copy of HCF repeat 2 was placed immediately carboxy-terminal to the Oct-1 DNA-binding POU domain. HCF repeat 2 was chosen because it matches the HCF repeat consensus described below at all but one position. The structure of the resulting Oct-1/rep2 chimeric protein is shown in Figure 4A. To differentiate this protein from endogenous Oct-1, the αHA epitope was included at the amino terminus.

The Oct-1/rep2 protein was expressed in 293 cells. Figure 4B shows a series of immunoblots of protein extracts prepared from transfected cells and probed with either the αHA-epitope antibody or with three sets of monoclonal antibodies specific for the amino-terminal (N$_t$), POU domain (POU), or carboxy-terminal (C$_t$) regions of Oct-1 (Lai and Herr 1992; Lai 1993). Expression of the parental “wild-type” epitope-tagged Oct-1 protein (wt) resulted in a major ~100-kD polypeptide recognized by the αHA-epitope antibody (cf. lanes 1 and 2). In contrast, expression of the HCF repeat-containing Oct-1/rep2 chimeras generated two novel polypeptides, a major ~110-kD product that probably corresponds to full-length Oct-1 with the additional HCF repeat 2, and a less abundant ~50-kD polypeptide (lane 3).

The 50-kD polypeptide is of the appropriate size for an amino-terminal Oct-1 fragment truncated at the inserted repeat (N$_t$+POU; see Fig. 4A). Consistent with this interpretation, the same sized polypeptide was detected by the N$_t$ [lane 7] and POU domain [lane 11] monoclonal antibodies but not the C$_t$ monoclonal antibodies [lane 15]. The Oct-1 C$_t$ antibodies identified a new smaller polypeptide [designated C$_t$; lane 15] that probably corresponds to a polypeptide extending from the site of insertion of HCF repeat 2 to the carboxyl terminus of Oct-1 [see Fig. 4A].

To confirm that cleavage of Oct-1/rep2 is specific to the HCF repeat sequence, we assayed a version of Oct-1 carrying an HCF repeat 2 sequence with the five alanine “KO” substitution mutation [Oct-1/rep2[KO]]. Expression of this protein again generated the 110-kD Oct-1/rep2 polypeptide but did not give rise to either the N$_t$+POU polypeptide [cf. lanes 3, 7, and 11 with lanes 4, 8, and 12] or the C$_t$ polypeptide [cf. lanes 15 and 16], consistent with complete inhibition of processing. These results indicate that HCF repeat 2 alone is sufficient to direct proteolytic cleavage of a heterologous protein. The inactivity of rep2[KO] demonstrates that recognition and cleavage of the inserted HCF repeat 2 sequence is sequence specific in this assay and is unlikely to result from a general alteration of the amino-terminal Oct-1 fragment.
from accidental formation of an exposed, protease-sensitive loop.

Multiple HCF repeats induce multiple cleavages of a heterologous protein

The fraction of precursor that was cleaved in the Oct-1/rep2 assay was low in comparison to transiently expressed HCF (cf. Figs. 3 and 4). We estimate that only a few percent of the total Oct-1/rep2 molecules were processed. This apparent inefficiency could result from (1) very high expression of the Oct-1/rep2 chimera saturating any machinery that may be required for processing; (2) rapid degradation of the Oct-1/rep2 N, + POU and C, cleavage products; or (3) inefficient processing of a single HCF repeat. To test the latter hypothesis, we compared the efficiency of Oct-1/rep2 processing to that of an Oct-1 molecule (Oct-1/rep123) carrying a segment of HCF containing the first three HCF repeats.

Figure 5A shows the structure of the Oct-1/HCF chimeras assayed, and Figure 5B shows an immunoblot of transfected 293 cell extracts probed with the aOct-1 POU domain monoclonal antibody. In addition to an expected larger precursor polypeptide, the Oct-1/rep123 chimera (lane 4) produced three prominent shorter products (indicated by solid arrows). The different sizes of these products correspond to the spacing between the HCF repeats in the inserted sequence, again indicating that specific cleavage is directed by each of the three HCF repeats. Each one of the three cleavage products, including the one that probably results from cleavage at HCF repeat 2 (fragment 2), was more than twice as abundant as the product generated by the isolated HCF repeat 2 in Oct-1/rep2 (cf. lanes 3 and 4). These results indicate that the presence of multiple repeats in a more natural context results in a significant increase in proteolytic processing and also that the low level of processed Oct-1/rep2 protein is not solely attributable to instability of the cleavage products.

The site of proteolytic cleavage lies within the HCF repeat

Figure 6B shows the sequence of the six near-perfect HCF repeats (1, 2, 3, 5, 6, and 8) and the two divergent repeats (4 and 7). To identify the sites of HCF cleavage in a population of native HCF fragments, we sequenced the amino termini of a highly purified preparation of HeLa cell HCF, containing both amino- and carboxy-terminal HCF fragments. The determined sequence is shown in Figure 6A. It corresponds to residues 11-19 of the six near-perfect HCF repeats [Fig. 6B], including one variable position (position 19) that can be either asparagine [N] or histidine [H] [Fig. 6A, see cycle 8]. These results indicate that cleavage occurs predominantly between the glutamic acid and threonine residues in the HCF repeat CE/TH sequence. Although we sequenced a pool of amino- and carboxy-terminal HCF fragments, we did not obtain a sequence that matches the amino terminus of the HCF ORF, suggesting that the amino terminus of the amino-terminal fragments is modified and is not susceptible to Edman degradation. Our proposed cleavage site matches the major one identified by Kristie et al. [1995] from amino-terminal sequence analysis of separated carboxy-terminal HCF fragments.

The HCF repeat is a highly specific proteolytic cleavage signal

To identify residues in the HCF repeat that are important for directing proteolytic cleavage, we constructed a
The HCF repeat signals proteolytic cleavage

Figure 6. Determination of the HCF repeat cleavage site and identification of functionally important residues in HCF repeat 2. [A] Partially purified human HeLa cell HCF was subjected to direct amino-terminal sequencing. The predominant amino acid residue determined in each of the first nine cycles is given in one-letter code. No specific residue could be determined in cycle 7, and in cycle 8 both asparagine and histidine were represented with similar abundance. The deduced cleavage site is indicated. [B] Alignment of human HCF repeats. The six near-perfect copies of the HCF repeat motif are grouped, with the two more degenerate copies [repeats 4 and 7] shown separately. Residues matching a consensus based on the most frequent residue at each position are stippled. At position 23, the amino acids valine and alanine are used with equal frequency but alanine was chosen for the consensus as this residue is preferred in the conserved repeats. The arrow marks the site of peptide bond cleavage between Glu-10 and Thr-11 as determined in [A]. [C] Summary of the relative extent of cleavage of single alanine substitution mutants of HCF repeat 2 assayed in the context of Oct-1I/rep2. The efficiency of specific proteolytic cleavage of each repeat derivative was estimated relative to Oct-1I/rep2 [wild type]. Mutants that showed a reduced but readily detectable level of activity were assigned an intermediate value. Residues Ala-20 and Ala-23 were not assayed. These values are based on between five and seven independent assays. The arrow indicates the predicted site of peptide bond cleavage between residues Glu-10 and Thr-11.

Panel of HCF repeat 2 amino-acid substitution mutants, in which each of the 24 nonalanine residues of this HCF repeat was changed in turn to alanine. These mutations were assayed in the context of the Oct-1I/rep2 chimera, and the relative extent of processing was determined by immunoblotting extracts of transfected 293 cells. The extents of cleavage were divided into three groups: wild type, not detectable, and intermediate levels. The results of multiple assays of each mutant are summarized in Figure 6C. They reveal a large 18-amino-acid segment of the HCF repeat [residues 7–24], within which most residues are shown to be important for HCF repeat-directed cleavage. Alanine substitution of flanking residues [residues 1–6 and 25 and 26] had little effect, suggesting that these positions are less critical for HCF repeat function. Although in this assay the importance of some conserved residues, for example Glu-13 and Gly-15, was not revealed, the extensive effects of the alanine substitutions suggest that the HCF repeat is highly conserved for the purpose of directing cleavage of HCF.

Typically, residues flanking one or both sides of a proteolytic cleavage site are critical for cleavage. We therefore assayed the effects of various substitutions of the Glu-10 and Thr-11 residues that border the deduced HCF repeat cleavage site [this study, Kristie et al. 1995]. Figure 7 shows the results of such an assay. Substitution of either residue with alanine [Fig. 7A, lanes 3,5] reduced cleavage to background levels [see also Fig. 6C], showing that both sides of the HCF repeat cleavage site are important for cleavage. The HCF repeat was also inactivated when Glu-10 was substituted for the more similar aspartic acid residue [cf. lanes 2 and 4], suggesting that the nature of this amino-terminal flanking residue is critically important. In contrast, substitution of the other flanking residue, Thr-11, with serine or phenylalanine had only intermediate effects on HCF repeat cleavage [cf. lanes 6 and 7 with lane 2], suggesting that the nature of this position is less critical than that of the amino-terminal flanking position. This result is consistent with other proteolytic cleavage sites, in which the residue amino terminal of the cleavage site is frequently more important than the one carboxy-terminal of the cleavage site [Polgar 1989; Thornberry et al. 1992; Welsh et al. 1993].
Proteolytic cleavage of Oct-1 is indicated with an open arrow. Full-length polypeptides that are recognized nonspecifically by the antibody to Oct-1 and Oct-ll/rep2 polypeptides are also indicated. Cellular extracts were prepared and the proteins assayed by immunoblotting with the aHA-epitope antibody (12CA5). The 50-kD polypeptide fragment corresponding to the N, POU body is shown indicating the proposed site of cleavage between Glu-4 and Thr-11. Individual substitution mutations of the two residues flanking the cleavage site are listed as well as the five alanine substitution mutation (KO) that includes this site.

**Figure 7.** Importance of the residues immediately flanking the HCF repeat cleavage site. [A] 293 cells were transfected with 5 μg of expression constructs for Oct-1 (lane 1), Oct-1/rep2 (lane 2), Oct-1/rep2(10A) (lane 3), Oct-1/rep2(10D) (lane 4), Oct-1/rep2(11A) (lane 5), Oct-1/rep2(11D) (lane 6), Oct-1/rep2(11S) (lane 7), and Oct-1/rep2(KO) (lane 8). Forty hours after transfection, whole-cell nuclear extracts were prepared and the proteins assayed by 10% SDS-polycrylamide gel electrophoresis followed by immunoblotting with the aHA-epitope antibody (12CA5). The 50-kD polypeptide fragment corresponds to the N, POU region of Oct-1 and indicated with an open arrow. Full-length Oct-1 and Oct-1/rep2 polypeptides are also indicated. Cellulose polypeptides that are recognized nonspecifically by the antibody are marked with an asterisk. Sizes of pre-stained protein markers (MW) are indicated. (B) The sequence of HCF repeat 2 is shown indicating the proposed site of cleavage between Glu-10 and Thr-11. Individual substitution mutations of the two residues flanking the cleavage site are listed as well as the five alanine substitution mutation (KO) that includes this site.

**An imperfect HCF repeat fails to induce proteolytic cleavage**

The alanine-scan mutagenesis described above predicts that even small variations in the HCF repeat sequence, such as those found in the two variant HCF repeats 4 and 7, will disrupt HCF repeat function. To test this hypothesis, we assayed the ability of HCF repeat 7 to direct cleavage of an Oct-1/rep7 chimera. We chose HCF repeat 7 because it is the least divergent of the two degenerate HCF repeats (see Fig. 6B). Figure 5B (lane 5) shows that the resulting Oct-1/rep7 polypeptide is expressed well in 293 cells but does not generate a specific cleavage product of the appropriate size (open arrow). This result suggests that the two variant HCF repeats 4 and 7 fail to direct proteolytic processing and further emphasizes the high degree of sequence specificity required for HCF repeat function.

**Proteolytic processing of HCF generates amino- and carboxy-terminal fragments that remain noncovalently associated**

We have shown previously that amino-terminal HCF fragments can be recovered by immunoprecipitation with antibodies raised against a carboxy-terminal HCF fragment, suggesting that amino- and carboxy-terminal fragments remain associated after HCF cleavage [Wilson et al. 1993b]. Those experiments, however, did not determine the nature of the association. To differentiate covalent from noncovalent association, we assayed the effect of protein denaturation on the recovery of amino- or carboxy-terminal HCF fragments using antibodies directed against the other half of the protein.

The results of such an experiment are shown in Figure 8A. A 293 cell extract containing transiently expressed epitope-tagged HCF was divided into two portions, one of which was denatured by boiling in the presence of SDS. The extracts were denatured under nonreducing conditions to preserve disulfide bonds. Recombinant HCF proteins were recovered from both the native (lanes 1, 2) and denatured (lanes 3, 4) extracts by immunoprecipitation with either the amino-terminal-specific epitope-tag antibody [αHCFN, lanes 1,3] or the carboxy-terminal-specific HCF antibody [αHCFC, lanes 2, 4]. The recovered proteins were resolved on a SDS-polycrylamide gel, and amino-terminal polypeptides were detected by immunoblotting with the αHCFN antibody (Fig. 8A, top).

Immunoprecipitation of the native extract with either the amino- or carboxy-terminal-specific antibody recovered similar sets of amino-terminal HCFN polypeptides (Fig. 8A, cf. lanes 1 and 2), consistent with efficient association of HCFN and HCFC polypeptides. Immunoprecipitation of the denatured extract (lanes 3, 4), yielded a different result. Denaturation had little, if any, effect on recovery of the amino-terminal HCFN polypeptides with the amino-terminal-specific αHCFN antibody (cf. lanes 1 and 3) but disrupted recovery of the smaller cleaved amino-terminal HCFN polypeptides with the carboxy-terminal-specific αHCFC antibody (cf. lanes 2 and 4). These results indicate that amino- and carboxy-terminal HCF polypeptides remain noncovalently associated after cleavage of the HCF precursor. Unlike the cleaved HCF fragments and consistent with its status as the full-length primary HCF translation product, recovery of HCF300 was not disrupted by denaturation (cf. lanes 2 and 4). This result shows that the failure to retrieve the smaller denatured HCF fragments is not attributable to denaturation of the epitope.

The bottom of Figure 8A shows the result of reprobing the same filter shown in the top with the carboxy-terminal-specific monoclonal antibody αHCF-M2 (see Fig. 1). This allows the sizes of the amino- and carboxy-terminal HCF fragments to be compared directly. The αHCFC monoclonal antibody identified a pattern of car-
Alternative mRNA splicing generates an HCF variant in which amino- and carboxy-terminal HCF fragments fail to associate

Previously, we have described a variant HCF cDNA, probably resulting from alternative mRNA splicing, that lacks sequences encoding HCF residues 382–450 [Wilson et al. 1993b; see Fig. 1]. Figure 8B shows that this 69-amino-acid deletion has no evident effect on HCF processing but disrupts subsequent association of the processed HCF fragments. Epitope-tagged HCF, either full length (HCF_{FL}) or lacking the alternative 69-amino-acid segment (HCF_{A382-450}), was expressed in 293 cells, and HCF fragments were recovered from native extracts by immunoprecipitation with the amino-terminal-specific aHCF_{N} epitope antibody (aHCF_{N}). Immune complexes were resolved on a 7% SDS–polyacrylamide gel and detected by immunoblotting with the amino-terminal-specific HCF, antibody (aHCF_{N}, lanes 1,2) or the aHCF-M2 monoclonal antibody (aHCF_{C}, lanes 3,4).

boxy-terminal HCF_{C} polypeptides very similar in number and size to the amino-terminal HCF_{N} polypeptides. This result explains why, when we originally determined the sequence of peptides obtained from individual bands of polyacrylamide gel-purified HeLa cell HCF [Wilson et al. 1993b], we obtained sequences derived from both the amino- and carboxy-terminal regions of the very large HCF coding sequence: the amino- HCF_{N} and carboxy-terminal HCF_{C} polypeptides comigrated in the individual bands that we sequenced.

Consistent with the results with the amino-terminal antibody [Fig. 8A, top], the carboxy-terminal antibody also generated very similar patterns irrespective of whether the native extract was first immunoprecipitated with the amino- or carboxy-terminal antibody [lanes 1',2']. Likewise, denaturation affected recovery of the smaller carboxy-terminal HCF_{C} fragments with the amino-terminal-specific aHCF_{N} antibody but not with the carboxy-terminal-specific aHCF_{C} antibody [lanes 3',4'], again consistent with noncovalent association of cleaved amino- and carboxy-terminal HCF polypeptides.

Alternative mRNA splicing generates an HCF variant in which amino- and carboxy-terminal HCF fragments fail to associate

Previously, we have described a variant HCF cDNA, probably resulting from alternative mRNA splicing, that lacks sequences encoding HCF residues 382–450 [Wilson et al. 1993b; see Fig. 1]. Figure 8B shows that this 69-amino-acid deletion has no evident effect on HCF processing but disrupts subsequent association of the processed HCF fragments. Epitope-tagged HCF, either full length (HCF_{FL}) or lacking the alternative 69-amino-acid segment (HCF_{A382-450}), was expressed in 293 cells, and HCF fragments were recovered from native extracts by immunoprecipitation with the amino-terminal-specific aHCF_{N} epitope antibody. The recovered HCF polypeptides were detected by immunoblotting with the amino-terminal aHCF_{N}, lanes 1,2) and carboxy-terminal aHCF_{C}, lanes 3,4) antibodies. When the same amino-terminal-specific HCF_{N} antibody was used for immunoprecipitation and for the immunoblot, the full-length HCF_{FL} and deleted HCF_{A382-450} proteins yielded similar patterns of amino-terminal HCF polypeptides, except that the HCF_{A382-450} polypeptides, which should be 69 amino acids shorter than the HCF_{FL} proteins, migrated more rapidly (cf. lanes 1 and 2 in this experiment the HCF_{300} precursor is not visible). The similar pattern of mature HCF fragments in the two samples suggests that the HCF_{A382-450} variant is processed normally.

However, analysis of the pattern of carboxy-terminal rHCF_{C} fragments recovered with the amino-terminal-specific aHCF_{N} antibody is very different for the HCF_{FL} and HCF_{A382-450} proteins. The HCF_{FL} protein yielded an expected pattern of full-length HCF_{300} and smaller HCF_{C} polypeptides attributable to association of processed amino- and carboxy-terminal HCF fragments [Fig. 8B, lane 3], whereas the shorter HCF_{A382-450} protein yielded only a slightly faster migrating form of HCF_{300}.

Figure 8. Noncovalent association of cleaved amino- and carboxy-terminal HCF fragments. (A) A whole-cell nuclear extract was prepared from 293 cells transfected with 5 μg of aHA epitope-tagged rHCF_{N} expression plasmid. Recombinant HCF polypeptides were recovered by immunoprecipitation either directly from aliquots of the native extract [lanes 1 and 2] or from aliquots that were first denatured by boiling in 2% SDS [lanes 3 and 4]. The proteins were immunoprecipitated using either a monoclonal antibody against the aHA epitope tag (aHCF_{N}, lanes 1,3) or the aHCF-H12 polyclonal antiserum (aHCF_{N}, lanes 2,4). The immune complexes were resolved on a 7% SDS–polyacrylamide gel and detected by immunoblotting with the aHA-epitope antibody (top, lanes 1–4). (Bottom lanes 1’–4’) The same blot after stripping and reprobing with aHCF-M2, which recognizes an epitope in the carboxy terminus of HCF [Wilson et al. 1995]. A minor 200-kD polypeptide in which the aHA and aHCF-M2 epitopes remain covalently linked is indicated with an asterisk. Molecular mass markers are given in kD. (B) An alternative splice variant prevents amino- and carboxy-terminal HCF polypeptide coassociation. 293 cells were transfected with 5 μg aT7-epitope-tagged expression plasmids encoding rHCF_{N} (lanes 2,4) and rHCF_{FL} (lanes 1,3). Recombinant HCF polypeptides were recovered from native extracts by immunoprecipitation with the aT7.tag antibody (arHCF_{N}). Immune complexes were resolved on a 7% SDS–polyacrylamide gel and detected by immunoblotting with the amino-terminal-specific aHCF_{N} antibody (lanes 1,2) and carboxy-terminal-specific (arHCF, antibody (lanes 3,4).
[Page 2454] Wilson et al.

A model for the maturation of human HCF

Figure 9 shows a simple schematic of HCF maturation. Much of our understanding of this process has been gleaned from the pulse–chase experiment shown in Figure 2. This experiment showed that the largest HCF polypeptide, HCF₃₀₀, is expressed first and is likely to be the primary HCF translation product. At first, HCF₃₀₀ was most abundant in the cytoplasmic fraction but, within about an hour after synthesis, became most abundant in the nuclear fraction. Thus, the majority of HCF appears to be transported to the nucleus as the HCF₃₀₀ precursor, which is then cleaved at the HCF repeats into smaller amino-terminal HCFₚ and carboxy-terminal HCFₙ polypeptides. For simplicity, in Figure 9 we show the association of HCFₚ and HCFₙ polypeptides derived from the same HCF₃₀₀ precursor, but it is possible that intermolecular exchange occurs.

The disappearance of nuclear HCF₃₀₀ is relatively slow, having an ~3-hr half-life in the transient expression assay shown in Figure 2. The transport of HCF₃₀₀ to the nucleus appears more rapid. Therefore, it is not yet evident whether HCF₃₀₀ is processed in the nucleus because that is where the HCF processing machinery is located or because it is simply transported to the nucleus more rapidly than it is processed. Because some processed HCF fragments are present in the cytoplasmic fraction up to 3 hr after HCF synthesis, the latter may be the case.

The accumulation of HCF₃₀₀ polypeptides in the nuclear fraction suggests that processing is not required for transport into the nucleus and is consistent with our ability to purify both HCF₃₀₀ and the smaller polypeptides from nuclear extracts (Wilson et al. 1993b). This observation suggests that HCF processing is not used to regulate extranuclear confinement, as has been documented for the NF-kB/Rel precursors p105 and p100 (Blank et al. 1991; Fan and Maniatis 1991; Mellits et al. 1993; Mercurio et al. 1993; Palombella et al. 1994; for review, see Liou and Baltimore 1993), and the precursor to the sterol-response factor SREBP-1 (Wang et al. 1994), both of which function as transcription factors.

The initial pattern of HCFₚ fragment accumulation in the nucleus suggests that most, if not all, six near-perfect HCF repeats are used with similar efficiency. This interpretation is substantiated by the pattern of cleaved polypeptide accumulation when three HCF repeats were inserted into Oct-1 (see Fig. 5). Subsequently, further proteolytic cleavage of the primary HCFₚ and HCFₙ fragments at the remaining HCF repeats results in “trimming,” and this leads to a general accumulation of the smaller HCFₚ and HCFₙ fragments. Indeed, in a survey of human cell lines and tissues, the smallest HCFₙ polypeptide was the predominant carboxy-terminal species (Wilson et al. 1995). Thus, as shown in Figure 9, HCF probably undergoes multiple rounds of HCF repeat-directed cleavage, with HCF repeats being selected in a largely stochastic manner. The destiny of the internal HCF fragments, released by cleavage at multiple HCF repeats, is unknown.

A surprisingly large recognition signal

The alanine scan study of HCF repeat 2 identified 11 residues that are important for directing cleavage, with an additional 6 residues that appear to play a less important role. With the exception of Thr-24, all of the resi-
The HCF repeat signals proteolytic cleavage

The only known biological function of HCF is in stabilization of the HSV VP16-induced complex. HCF processing, however, does not appear to be necessary for this function: Both HCFₐ and HCF₇ fragments are incorporated into the VP16-induced complex (Wilson et al. 1993b; Kristie et al. 1995), the HCFₗ₈ rep protein lacking the HCF repeats can associate normally with VP16, and, in the absence of the carboxyl terminus of HCF, a discrete amino-terminal domain of HCF can induce complex formation with VP16 and Oct-1 (A.C. Wilson and W. Herr, unpubl.). It is instead more likely that processing is required for a natural cellular function of HCF.

There are relatively few examples of intracellular proteins in which polypeptides generated by the processing of a common precursor remain associated with each other. The active form of the ICE protease functions as a tetramer composed of two p20/p10 heterodimers. These subunits are derived from a 45-kD proenzyme by removal of a 11.5-kD propeptide and a short internal linker peptide through intramolecular autocatalysis (for review, see Thornberry and Molineaux 1995). It is apparent from the crystal structure of the enzyme that the active sites are formed by the interface between p10 and p20 subunits of adjacent heterodimers (Walker et al. 1994; Wilson et al. 1994). Dissociation of the heterodimers prevents catalysis and, although not yet demonstrated in this instance, controlled coassociation may provide a means of regulation within the cell.

Another example of postcleavage coassociation is the two larger subunits of the general transcription factor TFIIB, which, like HCF, functions within the nucleus. In humans and insects these subunits are encoded by the amino- and carboxy-terminal ends of a single large ORF (DeJong and Roeder 1993; Ma et al. 1993; Yokomori et al. 1993). Although it is unclear whether these subunits arise through proteolytic processing of a larger precursor or by some other mechanism, both subunits remain associated with each other in conjunction with a third subunit encoded by a separate gene. Because none of the established in vitro assays for TFIIB function require subunit separation, it remains unclear why the two TFIIB subunits are generated from a single ORF (DeJong et al. 1995).

The tight association of the amino- and carboxy-terminal fragments after proteolytic cleavage remains one of the most puzzling features of HCF processing. This situation is reminiscent of the tetanus and botulinum toxins (for review, see Schiavo et al. 1994), and these bacterial proteins may provide a clue to the purpose of HCF processing. After release from the bacterial cell, these potent neurotoxins are produced by cleavage of an inactive 150-kD precursor polypeptide at a single site. The resulting 50- and 100-kD polypeptides remain associated by a single interchain disulfide bond that is essential for neurotoxicity. The 100-kD subunit is responsible for targeting the toxin to neuronal cells and also directs the insertion of the toxin into the membranes of endosome-like vesicles. It is thought that the reducing conditions of the endosomes are sufficient to break the interchain disulfide bond releasing the 50-kD subunit into

Why is HCF processed?

The association between processed HCF fragments is tight, and HCFₐ and HCF₇ polypeptides will remain associated even in the presence of 3 M urea (Kristie and Sharp 1993). The association is also efficient because the immunoprecipitation experiments presented here suggest that the majority of processed fragments remain associated. In Figure 8A, for instance, recovery of the major amino-terminal polypeptides was as efficient using the carboxy-terminal aHCF antisera as with the amino-terminal αHA-epitope antibody (cf. lanes 1 and 2), suggesting that almost every amino-terminal fragment encoded by the full-length HCF cDNA is paired with a carboxy-terminal fragment. Thus, if the cleavage products remain associated with one another, why is HCF cleaved in the first place?

dues that we identified as important are perfectly conserved in the six near-perfect HCF repeats. Further emphasizing the importance of the HCF repeat sequence, the corresponding six HCF repeats in hamster HCF differ from the human sequence at only one position in one repeat (GenBank accession no. D45419; H. Goto and T. Nishimoto, pers. comm.). When compared to other known protease recognition sites this large number of critical residues is extraordinary. We have not identified convincing matches to the HCF repeat in the sequence database arguing that the HCF repeat is not a frequently used proteolytic cleavage signal. Indeed, it may be dedicated to HCF maturation.

The deduced location of HCF repeat cleavage is also unusual because the majority of residues important for cleavage lie carboxy-terminal of the cleaved bond. Usually, the preponderance of important residues lie amino-terminal of the cleavage site. Virus-encoded proteases, the cellular interleukin-1β-converting enzyme (ICE), and subtilisin-like proteases are among the best characterized endoproteases (for review, see Kräusslich and Wimmer 1988; Barr 1991; Thornberry and Molineaux 1995). Mutagenesis studies of natural and synthetic substrates have shown that these enzymes generally require specific residues at only a few positions on either side of the scissile bond for cleavage. In general, no more than four specific residues are required on the amino-terminal side (positions P1–P4) of the cleavage site, and even fewer residues are required on the carboxy-terminal-side (P' positions) (Polgar 1989; Thornberry et al. 1992; Welsh et al. 1993).

One possibility is that the HCF repeat consists of more than one functional element. For instance, there may be a traditional shorter protease recognition site surrounding the proposed cleavage site between Glu-10 and Thr-11, and one or more accessory sites spanning the threonine-rich carboxy-terminal end of the repeat. All eight threonine residues in HCF repeat 2 are functionally important, and it is possible that some serve as phosphorylation substrates. Phosphorylation or another post-translational modification might allow regulation of processing or increase the specificity of HCF repeat recognition.

The association between processed HCF fragments is tight, and HCFₐ and HCF₇ polypeptides will remain associated even in the presence of 3 M urea (Kristie and Sharp 1993). The association is also efficient because the immunoprecipitation experiments presented here suggest that the majority of processed fragments remain associated. In Figure 8A, for instance, recovery of the major amino-terminal polypeptides was as efficient using the carboxy-terminal aHCF antisera as with the amino-terminal αHA-epitope antibody (cf. lanes 1 and 2), suggesting that almost every amino-terminal fragment encoded by the full-length HCF cDNA is paired with a carboxy-terminal fragment. Thus, if the cleavage products remain associated with one another, why is HCF cleaved in the first place?
the cytosol. Once released, the small subunit acts as a zinc endopeptidase, degrading proteins of the neuroexocytotyosis apparatus. For efficient function these toxins require the cooperation of several distinct elements. Protopolytic cleavage, subsequent coassociation, and regulated release presents a mechanism for coordinating these functions.

We favor a similar role for processing and coassociation in HCF. Separate regions of HCF may perform different functions. These regions may need to function in unison but may also need to dissociate in a regulated manner. The alternative splice variant of HCF (see Fig. 8B) suggests that HCF$_C$- and HCF$_C$-fragment association can be regulated, in this case, by removal of a small region of the HCF$_C$ fragments. Perhaps, HCF$_N$ and HCF$_C$ fragment association is a dynamic and regulated process in vivo. For example, one of the HCF$_N$ fragments recovered by immunoprecipitation with a HCF$_C$-specific antibody in Figure 8A is not retrieved by immunoprecipitation with an HCF$_C$-specific antibody (cf. lanes 1 and 2, solid arrowhead); perhaps, this particular HCF$_N$ polypeptide is modified so that it fails to associate with HCF$_C$ fragments.

This dynamic view of HCF$_N$ and HCF$_C$ polypeptide association suggests that mature fragments are stable both in association with each other and when on their own. Indeed, HCF$_N$ and HCF$_C$ fragments may be expressed as a single precursor protein primarily to ensure stoichiometric levels of HCF$_N$ and HCF$_C$ polypeptides in the cell, as has been proposed for a number of viral polyproteins [see Kräusslich and Wimmer 1988]. Protopolytic cleavage, directed by the HCF repeats, then permits the amino- and carboxy-terminal domains of HCF to function either together or separately.

Materials and methods

Expression constructs

The epitope-tagged, full-length, HCF expression construct pCGNHCF$_{9E10}$ [designated previously as pCGNHCF(2-2035)9E10] has been described previously [Wilson et al. 1993b]. The expression plasmid pCGTHCFFL is equivalent to pCGNHCF$_{9E10}$ except that the amino-terminal HA epitope tag was replaced by the sequence MasMTGGQQMG. This sequence corresponds to the first 11 residues of the bacteriophage T7 gene 10 capsid protein and is recognized by the monoclonal antibody aHCF-H12 [Wilson et al. 1993b].

For metabolic labeling cells were transfected by electroporation using a Bio-Rad Gene pulser with Extendor, set at 200 V and 960 μF. Whole-cell nuclear extracts were prepared after 40 hr by lysing cells in medium salt lysis buffer [250 mM KCl, 10 mM HEPEs-KOH (pH 7.9), 5% glycerol, 0.1% NP-40, 0.2 mM EDTA, 0.5 mM PMSF, 0.2 mM sodium vanadate, 50 μM sodium fluoride, 1 mM DTT, 1 mM benzamidine, 10 μg/ml leupeptin, 10 μg/ml aprotinin]. Lysates were incubated at 4°C for 30 min to allow elution of nuclear proteins before removal of the nuclei and other debris by centrifugation.

Immunoprecipitation and immunoblotting

Immunoprecipitation was performed using both native and denatured extracts. HCF polypeptides were recovered from denatured extracts as follows: Samples were denatured prior to immunoprecipitation by the addition of one-tenth volume of 20% SDS and incubated in a boiling water bath for 10 min. The SDS was diluted by rapid addition of 20 volumes of cold (4°C) RIPA buffer lacking SDS [150 mM NaCl, 50 mM Tris-HCl (pH 8.0), 1% NP-40, and 0.5% DOC]. Typically, the equivalent of one-fourth of a 10-cm plate of transfected 293 cells was used per immunoprecipitation. Extracts were incubated with antibodies—10 μl of aHA [12CA5] mouse ascites fluid or 40 μl of rabbit aHCF-H12 serum (or HCF, Wilson et al. 1993b)—at 4°C for 1 hr with rotation. Immune complexes were recovered with protein G-agarose beads (GIBCO) and washed four times with 1 ml of cold RIPA buffer without SDS. For immunoprecipitation under native conditions, antibodies were added directly to whole-cell nuclear extracts. Immunoblotting was performed by semidry transfer as described in Wilson et al. (1993b) and detected by Enhanced Chemiluminescence [Amersham]. The aHA- and aHCF-M2 (Wilson et al. 1995) monoclonal antibodies were used at a 1:5000 dilution. The aT7.tag monoclonal antibody was diluted 1:10,000. Anti-Oct-1 monoclonal antibody ascites [YL8 and YL15] were diluted 1:1000, and hybridoma supernatants [YL18 and YL36] were diluted 1:100. For sequential probing with a different antibody, blots were stripped of the first antibody by incubation in 100 mM 2-mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl (pH 6.7) at 50°C for 30 min. Blots were washed extensively in a large volume of Tritis-buffered saline [20 mM Tris-HCl at pH 7.6, 137 mM NaCl] to remove SDS, before rebloking and reprobing.

Metabolic labeling

For metabolic labeling cells were transfected by electroporation and pooled prior to plating to normalize for variations in transfection efficiency. Twenty hours after transfection, the cells were washed in phosphate-buffered saline [PBS] and incubated at 37°C for 30 min in 2 ml of prewarmed DMEM lacking serum but with 0.4 μCi of Trans35S-label [ICN Pharmaceuticals]. After the labeling period, the medium was removed and the cells were washed with 10 ml of prewarmed unlabeled media containing 10% calf serum. At appropriate times, the cells were collected and washed with PBS. Cells were lysed by gentle mixing for 1 min at 4°C in 1 ml of low salt lysis buffer [with 10 mM KCl], and the nuclei pelleted by microcentrifugation for 30 sec. The superna-
tact constituted the cytoplasmic fraction. The isolated nuclei were resuspended in 1 ml of high salt lysis buffer (250 mM KCl) and incubated for 30 min at 4°C. The nuclei were removed by centrifugation and the supernatant taken as the nuclear fraction.

Recombinant HCF polypeptides were recovered from heat- and SDS-denatured extracts by immunoprecipitation with αHA-epitope antibody. Before addition of the specific antibody, extracts were precleared with normal rabbit serum and fixed Staphylococcus aureus Cowan I (Zymed Laboratories). Immune complexes were resolved by 7% SDS–polyacrylamide gel electrophoresis, and visualized by fluorography using the enhancing agent Amplify (Amersham). Radioactivity was measured using a Fuji BAS1000 PhosphorImager.

Protein sequencing

HCF was purified from HeLa cell nuclear extracts by wheat germ agglutinin affinity and Mono-S chromatography as described previously [Wilson et al. 1993b]. This material was diluted to 100 mM KCl, 20 mM HEPES-KOH (pH 7.9), 0.5 mM EDTA, 20% glycerol, 1 mM DTT, 0.1% LDAO, and fractionated on Q–Sepharose resin, eluting with a 100–175 mM KCl step. Approximately 10 pmol of this material was subjected to amino-terminal sequence analysis on an ABI 477A with a 120A analyzer.

Acknowledgments

We thank Jon Tupy for technical assistance; Mike Brasseur for amino-terminal sequence determination; Hiroshi Goto and Takeharu Nishimoto for communicating the hamster HCF cDNA sequence, Ruyui Kobayashi for helpful discussions; Jim Duffy, Phil Renna, and Mike Ockler for artwork. We also thank Kim Gavin, Paul Kaufman, Kelly LaMarco, Yuri Lazebnik, Steve McKnight, and Robert Tjian for critical comments on the manuscript. A.C.W. was supported by a Damon Runyon–Walter Winchell Cancer Research Fund Fellowship [DRG-1073]. These studies were supported by U.S. Public Health Service grant CA-13106 from the National Cancer Institute and by Tularik, Inc. The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Barr, P.J. 1991. Mammalian subtilisins: The long-sought dibasic processing endopeptidases. Cell 66: 1–3.
Blank, V., P. Kourilsky, and A. Israel. 1991. Cytoplasmic retention, DNA binding and processing of the NF-κB p50 precursor are controlled by a small region in its C-terminus. EMBO J. 10: 4159–4167.
Dejong, J. and R.G. Roeder. 1993. A single cDNA, hTFIIA/α, encodes both the p35 and p19 subunits of human TFIIA. Genes & Dev. 7: 2220–2234.
Dejong, J., R. Bernstein, and R.G. Roeder. 1995. Human general transcription factor TFIIA: Characterization of a cDNA encoding the small subunit and requirement for basal and activated transcription. Proc. Natl. Acad. Sci. 92: 3313–3317.
Fan, C.-M. and T. Maniatis. 1991. Generation of p50 subunit of NF-κB by processing of p105 through an ATP-dependent pathway. Nature 354: 395–398.
Gerster, T. and R.G. Roeder. 1988. A herpesvirus trans-activating protein interacts with transcription factor OTF-1 and other cellular proteins. Proc. Natl. Acad. Sci. 85: 6347–6351.
Hunter, T. 1995. Protein kinases and phosphatases: The Yin and Yang of protein phosphorylation and signaling. Cell 80: 225–236.
Jackson, C.M. and Y. Nemerson. 1980. Blood coagulation. Annu. Rev. Biochem. 49: 765–811.
Jackson, S.P. and R. Tjian. 1988. O-glycosylation of eukaryotic transcription factors: Implications for mechanisms of transcriptional regulation. Cell 55: 125–133.
Katan, M., A. Haigh, C.P. Verrizer, P.C. van der Vliet, and P. O'Hare. 1990. Characterization of a cellular factor which interacts functionally with Oct-1 in the assembly of a multicomponent transcription complex. Nucleic Acids Res. 18: 6871–6880.
Kräusslich, H.-G., and E. Wimmer. 1988. Viral proteases. Annu. Rev. Biochem. 57: 701–754.
Kristie, T.M. and P.A. Sharp. 1993. Purification of the cellular C1 factor required for the stable recognition of the Oct-1 homeodomain by the herpes simplex virus α-trans-induction factor (VP16). J. Biol. Chem. 268: 6535–6544.
Kristie, T.M., J.H. LeBonne, S.H. Wang, and P.A. Sharp. 1989. The octamer-binding proteins form multi-protein-DNA complexes with the HSV αTIF regulatory protein. EMBO J. 8: 4229–4238.
Kristie, T.M., J.L. Pomerantz, T.C. Twomey, S.A. Parent, and P.A. Sharp. 1995. The cellular C1 factor of the herpes simplex virus enhancer complex is a family of polypeptides. J. Biol. Chem. 270: 4387–4394.
Lai, J.-S. 1993. “The herpes simplex virus VP16-induced complex: transcriptional regulation through selective protein–protein association.” Ph.D. thesis, State University of New York-Stony Brook, Stony Brook, New York.
Lai, J.-S. and W. Herr. 1992. Ethidium bromide provides a simple tool for identifying genuine DNA-independent protein associations. Proc. Natl. Acad. Sci. 89: 6958–6962.
Liou, H.-C. and D. Baltimore. 1993. Regulation of the NF-κB bel transcription factor and IκB inhibitor system. Curr. Opin. Cell Biol. 5: 477–487.
Ma, D., H. Watanabe, F. Mermelstein, A. Admon, K. Oguri, X. Sun, T. Wada, T. Imai, T. Shiroya, D. Reinberg, and H. Handa. 1993. Isolation of a cDNA encoding the largest subunit of TFIIA reveals functions important for activated transcription. Genes & Dev. 7: 2246–2257.
Mellits, K.H., R.T. Hay, and S. Goodbourn. 1993. Proteolytic degradation of MAD3 (IκBα) and enhanced processing of the NF-κB precursor p105 are obligatory steps in the activation of NF-κB. Nucleic Acids Res. 21: 5059–5066.
Mercurio, F., J.A. DiDonato, C. Rosette, and M. Karin. 1993. p105 and p98 precursor proteins play an active role in NF-κB-mediated signal transduction. Genes & Dev. 7: 705–718.
O’Hare, P. 1993. The virion transactivator of herpes simplex virus. Semin. Virol. 4: 145–155.
Palombella, V.J., O.J. Rando, A.L. Goldberg, and T. Maniatis. 1994. The ubiquitin-proteasome pathway is required for processing the NF-κB1 precursor protein and the activation of NF-κB. Cell 78: 773–785.
Polgar, L. 1989. Mechanisms of protease action. CRC Press, Boca Raton, FL.
Schiavo, G., O. Rossetto, and C. Montecucco. 1994. Clostridial neurotoxins as tools to investigate the molecular events of neurotransmitter release. Semin. Cell Biol. 5: 221–229.
Sturm, R.A., G. Das, and W. Herr. 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes & Dev. 2: 1582–1599.
Tanaka, M. and W. Herr. 1990 Differential transcriptional activation by Oct-1 and Oct-2: Interdependent activation domains induce Oct-2 phosphorylation. Cell 60: 375–386.

The HCF repeat signals proteolytic cleavage

GENES & DEVELOPMENT 2457
Wilson et al.

Tanaka, M., U. Grossniklaus, W. Herr, and N. Hernandez. 1988. Activation of the U2 snRNA promoter by the octamer motif defines a new class of RNA polymerase II enhancer elements. *Genes & Dev.* 2: 1764–1778.

Thompson, C.C. and S. L. McKnight. 1992. Anatomy of an enhancer. *Trends Genet.* 8: 232–236.

Thornberry, N.A. and S.M. Molineaux. 1995. Interleukin-1β converting enzyme: A novel cysteine protease required for IL-1β production and implicated in programmed cell death. *Protein Sci.* 4: 3–12.

Thornberry, N.A., H.G. Bull, J.R. Calaycay, K.T. Chapman, A.D. Howard, M.J. Kostura, D.K. Miller, S.M. Molineaux, J.R. Weidner, J. Aunins, K.O. Elliston, J.M. Ayala, F.J. Casano, J. Chin, G.J.F. Ding, L.A. Egger, E.P. Gaffney, G. Limjucuo, O.C. Palyha, S.M. Raju, A.M. Rolando, J.P. Salley, T.T. Yamin, T.D. Lee, J.E. Shively, M. MacCoss, R.A. Mumford, J.A. Schmidt, and M.I. Tocci. 1992. A novel heterodimeric cysteine protease is required for interleukin-1β processing in monocytes. *Nature* 356: 768–774.

Walker, N.P.C., R.V. Talanian, K.D. Brady, L.C. Dang, N.J. Bump, C.R. Ferenz, S. Franklin, T. Ghayur, M.C. Hackett, L.D. Hammill, L. Herzog, M. Hugunin, W. Houy, J.A. Mankovitch, L. McGuiness, E. Orlewicz, M. Paskind, C.A. Pratt, P. Reis, A. Summani, M. Terranova, J.P. Welch, L. Xiong, A. Miller, D.E. Tracey, R. Kamen, and W.W. Wong. 1994. Crystal structure of the cysteine protease interleukin-1β-converting enzyme: A (p20/p10), homodimer. *Cell* 78: 343–352.

Wang, X., R. Sato, M.S. Brown, X. Hua, and J.L. Goldstein. 1994. SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis. *Cell* 77: 53–62.

Welch, A.R., L.M. McNally, M.R.T. Hall, and W. Gibson. 1993. Herpes protease: Site-directed mutagenesis used to study maturational, release, and inactivation cleavage sites of precursor and to identify a possible catalytic site serine and histidine. *J. Virol.* 67: 7360–7372.

Wilson, A.C., M.A. Cleary, J.-S. Lai, K. LaMarco, M.G. Peterson, and W. Herr. 1993a. Combinatorial control of transcription: The herpes simplex virus VP16-induced complex. *Cold Spring Harbor Symp. Quant. Biol.* 58: 167–178.

Wilson, A.C., K. LaMarco, M.G. Peterson, and W. Herr. 1993b. The VP16 accessory protein HCF is a family of polypeptides processed from a large precursor protein. *Cell* 74: 115–125.

Wilson, A.C., J.E. Parrish, H.F. Massa, D.L. Nelson, B.J. Trask, and W. Herr. 1995. The gene encoding the VP16-accessory protein HCF (HCFC1) resides in human Xq28 and is highly expressed in fetal tissues and the adult kidney. *Genomics* 25: 462–468.

Wilson, K.P., J.F. Black, J.A. Thomson, E.E. Kim, J.P. Griffith, M.A. Navia, M.A. Murcko, S.P. Chambers, R.A. Aldape, S.A. Raybuck, and D.J. Livingston. 1994. Structure and mechanism of interleukin-1β converting enzyme. *Nature* 370: 270–275.

Xiao, P. and J.P. Capone. 1990 A cellular factor binds to the herpes simplex virus type 1 transactivator Vmw65 and is required for Vmw65-dependent protein-DNA complex assembly with Oct-1. *Mol. Cell. Biol.* 10: 4974–4977.

Yokomori, K., A. Admon, J.A. Goodrich, J.L. Chen, and R. Tjian. 1993. *Drosophila* TFIIA-L is processed into two subunits that are associated with the TBP/TAF complex. *Genes & Dev.* 7: 2235–2245.
The HCF repeat is an unusual proteolytic cleavage signal.

A C Wilson, M G Peterson and W Herr

*Genes Dev.* 1995, 9:
Access the most recent version at doi:10.1101/gad.9.20.2445

References

This article cites 37 articles, 14 of which can be accessed free at:
[http://genesdev.cshlp.org/content/9/20/2445.full.html#ref-list-1](http://genesdev.cshlp.org/content/9/20/2445.full.html#ref-list-1)

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here.](#)