JOURNAL OF STRUCTURAL BIOLOGY **122,** 76–85 (1998)
ARTICLE NO. SB983957

Small Proline-Rich Proteins Are Cross-Bridging Proteins in the Cornified Cell Envelopes of Stratified Squamous Epithelia[1]

Peter M. Steinert, Eleonora Candi, Tonja Kartasova, and Lyuben Marekov

Laboratory of Skin Biology, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland 20892-2752

Received November 3, 1997

The cornified cell envelope (CE) is a specialized structure which contributes barrier function to stratified squamous epithelial cells. It is composed of an amalgam of several structural proteins that are rendered insoluble by isopeptide bond crosslinking by transglutaminases. One set of the structural proteins present in CEs of most such epithelia are the small proline rich (SPR) proteins, which are a family of about 12 related structural proteins. We have recovered a large number of peptides containing isopeptide crosslinks, including 236 involving SPR proteins, following proteolysis of CEs isolated from foreskin epidermal tissue and cultured epidermal keratinocytes. Analysis of this database has provided novel information on their function. First, we found that SPRs became crosslinked to many other structural proteins within the CE. Second, multiple glutamine and lysine residues located only on the amino- and carboxy-termini of the SPR proteins were involved in crosslinking, so that the two ends are functionally equivalent. Third, the SPRs functioned as cross-bridging proteins, by directly adjoining other CE structural proteins. In the specialized case of the epidermal CE, the SPRs cross-bridged between loricrin. In cultured keratinocytes which make little loricrin and serve as a model for internal stratified squamous epithelia, the SPRs formed extensive cross-bridges among themselves. Thus SPRs are ubiquitous cross-bridging proteins whose differential expression patterns apparently reflect specific barrier requirements of different epithelia.

Key Words: **barrier function; cornified cell envelope; epidermis; involucrin; loricrin; transglutaminases.**

INTRODUCTION

Epithelial barrier function is essential for the life of mammals. This is established primarily by an insoluble protein layer located on the inner surface of the cell periphery in terminally differentiating stratified squamous epithelial cells, termed the cornified cell envelope (CE) (Hohl, 1990; Reichert *et al*., 1993; Simon, 1994). In highly specialized "dry" epithelial tissues such as the epidermis and hair fiber cuticle, a layer of lipids is also attached to the protein CE to provide water barrier function (Downing *et al*., 1993; Elias, 1996; Swartzendruber *et al*., 1987).

The protein portion of the CE is made insoluble by crosslinking of the constituent proteins through disulfide bonds, and N^{ϵ}-(γ-glutamyl)lysine or bis(γ-glutamyl)spermidine isopeptide bonds formed by the action of transglutaminases (TGases) (Greenberg *et al*., 1991; Martinet *et al*., 1988; Reichert *et al*., 1993). At least three different enzymes, TGases 1, 2, and 3, are involved in CE formation, and they appear to have complementary or overlapping roles in the crosslinking of the several structural proteins (Candi *et al*., 1995; Tarcsa *et al*., 1997). For example, the essential role of the TGase 1 enzyme has been confirmed by the fact that mutations in the *TGM1* gene encoding it cause the serious life-threatening disease lamellar ichthyosis (Huber *et al*., 1995; Russell *et al*., 1995). Several proteins, including cystatin α, desmoplakin, elafin, envoplakin, filaggrin, involucrin, five different keratin chains, loricrin, and multiple individual members of the small proline-rich (SPR) family, are now known to be isopeptide crosslinked components of the human epidermal CE (Hohl *et al*., 1991; Steinert and Marekov, 1995, 1997). Other calcium binding and desmosomal proteins may be components of CEs (Robinson *et al*., 1997), but their mode of covalent attachment is not yet known. In all cases of the epidermis so far examined in detail, loricrin is the major component, accounting for as much as 80% of the total CE protein mass (Steven and Steinert, 1994; Steinert and Marekov, 1995, 1997; Yoneda and Steinert (1993). However, it is not expressed in most internal epithelia (Hohl *et al*., 1993). Similarly, the amounts of the

[1]This is a solicited contribution from the Second Alpbach Workshop on Coiled-Coil and Fibrous Proteins, 1997.

1047-8477/98

SPR proteins vary from near zero in interfollicular epidermis to about 30% in certain epithelia or cultured keratinocytes (Hohl *et al*., 1995; Steven and Steinert, 1994; Jarnik *et al*., 1996).

While the role and function of loricrin in the epidermis has been explored in some detail (Hohl *et al*., 1991; Candi *et al*., 1995; Steinert and Marekov, 1995, 1997), relatively less is known about the function of the SPR proteins. SPRs consist of a heterogeneous population of proteins that are nevertheless built according to a common theme of amino- and carboxy-terminal domains rich in Lys and Gln residues, flanking a central domain of peptide repeats that are enriched in Pro residues (Gibbs *et al*., 1993; Hohl *et al*., 1995; Kartasova *et al*., 1988, 1996; Marvin *et al*., 1991). In mouse, rabbit, and human, the SPR3 class consists of one protein and its central domain contains about 23 conserved eight-residue peptide repeats. The SPR1 class consists of two members in mammals, which have Lys- and Gln-rich termini which are homologous to those of SPR3, but their central domains are composed of from as few as 6 (human), to 12 (rabbit), or 13 (SPR1a) or 14 (SPR1b) (mouse) eight-residue repeats rich in prolines of sequence rather different from those of SPR3. The SPR2 class consists of about eight different proteins in the mammalian species examined so far and also have Lys- and Gln-rich termini which are generally homologous to the other SPRs, but their central domains contain a different nine-residue motif very rich in prolines that is repeated 3.5–9 times. In all approximately 12 SPR proteins, the central peptide repeats contain multiple Gln and Lys residues.

Individual SPRs are differentially expressed in varying amounts in virtually all stratified squamous epithelia (An *et al*., 1993; Gibbs *et al*., 1993; Hohl *et al*., 1995; Kartasova *et al*., 1988, 1996). Most epithelial cell types express only a small subset of proteins, and interestingly, expression is often variable between different locations within the same epithelium, notably in different body locations of the epidermis. As examples, the SPR1a protein is expressed in epidermis, cultured epidermal keratinocytes, and many other internal epithelial cell types; SPR1b expression is limited to the foreskin, but is coexpressed with SPR1a in cultured keratinocytes. Likewise, a few members of the SPR2 class are expressed in epidermis, but others are expressed more abundantly in tissues such as oral epithelia and vagina. SPR3 expression is highest in oral epithelia, esophagus, and rodent forestomach. In contrast, a much larger selection of SPR proteins is expressed in UV-damaged, hyperproliferative, malignant, or chemically treated epithelia (An *et al*., 1993; Austin *et al*., 1996; Fujimoto *et al*., 1997, Hohl *et al*., 1995, Kartasova *et al*., 1988; Owens *et al*., 1996). However, the molecular and functional reasons or consequences for these differential expression patterns are not yet clear.

We have analyzed and sequenced peptides derived from CEs of foreskin epidermis which contains large amounts of loricrin but few SPRs (Steinert and Marekov, 1995, 1997), as well as from cultured epidermal keratinocytes which contain very little loricrin but much more SPRs and thereby serve as a model system for other internal epithelia (PMS and LNM, unpublished). In addition to the several other proteins listed above, these data contained many occurrences of SPR sequences which we have examined in detail in this paper. Our data provide novel insights into how the SPR proteins function as cross-bridging proteins in the CEs.

MATERIALS AND METHODS

Sequencing of CE peptides. Experimental details for the isolation, separation, characterization, and sequencing of peptides containing isopeptide crosslinks have been described for human foreskin cornified tissue (Steinert and Marekov, 1995), immature foreskin epidermis, and saponified cornified tissue (Steinert and Marekov, 1997).

Similarly, we have characterized peptides from CEs produced by 3- and 7-day cultured normal human epidermal keratinocytes grown in serum-free submerged cultures in the presence of 1.2 mM Ca^{2+}. Briefly, the keratinocytes (Clonetics) were plated onto collagen-coated dishes at a density of 5×10^3 cells/cm^2 in KGM medium, according to the manufacturer's recommendations, in low calcium medium (0.08 mM Ca^{2+}) for about 3 days to achieve confluency, and transferred to high calcium medium (1.2 mM Ca^{2+}) in the presence of 25 μg/ml of the calcium ionophore A23187 (Calbiochem) for 3 to 7 days to achieve an advanced degree of terminal differentiation in submerged cell culture (Yuspa *et al*., 1989; Kim *et al*., 1995). Attached cells and sloughed cells from the 7-day cultures were lysed, harvested by dissolution into a buffer of 8 M urea, 50 mM Tris–HCl (pH 8.4), 1 mM EDTA, and pelleted. The pellets were then extracted with exhaustive boiling and sonication in SDS buffer as described previously (Steinert, 1995; Steinert and Marekov, 1995). The recovered CE fragments were subjected to proteolysis with trypsin and proteinase K for up to 4 h, and the resulting peptides were resolved by HPLC and recovered for protein sequencing (P.M.S. and L.N.M., unpublished data). A total of 308 peptides were sequenced which contained one or more isopeptide crosslinks, yielding 770 "peptide branches." Detailed characterization of the proteins discovered in these experiments will be reported elsewhere. However, 180 of the peptide branches involved crosslinks with SPR proteins, which are examined in detail in this paper.

RESULTS

SPRs Are Cross-Linked by Isopeptide Bonds to Many Other CE Proteins

We have characterized more than 1600 peptides containing one or more isopeptide crosslinks from CEs isolated from intact human foreskin epidermis or cultured human epidermal keratinocytes. The

majority of these peptides involved loricrin (Steinert and Marekov, 1995, 1997). Other very abundant peptides involved involucrin (Steinert and Marekov, 1997) or keratins (Candi *et al*., 1998). A fourth set included 236 peptides which contained one or more SPR protein. The SPR proteins constitute a total of only about 5% (molar basis) of the protein of epidermal CEs (56 peptides) (Steinert and Marekov, 1997), but constituted 29% (3 day) or 21% (7 day) of the total protein of cultured keratinocyte CEs (182 peptides) (P.M.S. and L.N.M., unpublished). These recovered amounts are consistent with the amounts predicted from mathematical modeling of amino acid compositions (Steinert, 1995; Steven and Steinert, 1994).

Table I lists the Lys or Gln residues of the protein partners to which either the SPR1 (142 occurrences) or SPR2 (94 occurrences) proteins were found to be crosslinked. We have not recovered any peptides unambiguously derived from the SPR3 protein, perhaps since it is poorly expressed in foreskin epidermis and not at all in cultured keratinocytes, although a few listed as SPR1 may have arisen from SPR3 since their carboxy-termini are identical (Steinert and Marekov, 1995). The list contains representatives of all of the proteins known to date to be involved in epidermal CEs, with the single exception of filaggrin. In the cases of the keratins and desmoplakin, for example, there was a very high degree of specificity of the Lys residue(s) used in crosslinking to the SPRs. However, in involucrin and loricrin, multiple residue Lys and Gln positions were utilized.

Furthermore, many different Lys and Gln residues of the SPR1 (Table II) and SPR2 (Table III) proteins were used in crosslinks with these various protein partners. Moreover, in both SPR1 and SPR2, only those Lys and Gln residues located on the amino- and carboxy-termini were involved in crosslinks. In the case of SPR1a/b proteins, none of the 15 Lys and Gln residues of the central peptide repeats was used. Similarly, for SPR2, none of the 7 (shortest member) to 20 (longest member) central Lys and Gln residues was used. In both protein classes, the amino- or carboxy-termini were approximately equally used. That is, the two ends of the SPRs were functionally equivalent in crosslinking. However, most Gln crosslinking sites were located on the amino-terminus, whereas most Lys sites were located on the carboxy-terminus, and indeed, the most frequently used residue was the terminal Lys residue.

SPRs Function as Cross-Bridging Proteins in CEs

While some of the peptides recovered in the sequencing experiments involved a single isopeptide crosslink between an SPR protein and only one other protein partner, most contained two or more

TABLE I

SPRs are Crosslinked *in Vivo* to Many Protein Partners

	SPR1		SPR2	
Protein partner	Number	Residue of partner	Number	Residue of partner
Cystatin α	1	Lys46		
Desmoplakin	9	Lys1659	2	Lys1659
	5	Lys1661	1	Lys1661
Elafin	1	Gln2	1	Gln2
Envoplakin	7	Gln1970	4	Gln1970
	7	Gln1973	3	Gln1973
Involucrin	2	Gln308	2	Gln308
	1	Gln309	1	Gln309
	6	Gln368	5	Gln368
	2	Gln369	3	Gln369
	10	Gln425	9	Gln425
	6	Gln426	7	Gln426
	9	Gln455	7	Gln455
	3	Gln456	4	Gln456
Keratin 1	5	Lys73	4	Lys73
Keratin 2e	1	Lys69	1	Lys69
Keratin 5	8	Lys71	5	Lys71
Keratin 6	1	Lys68		
Loricrin	1	Gln3	1	Gln3
	1	Lys4	1	Lys4
	1	Lys5		
	2	Lys88	1	Lys88
	1	Gln153		
	1	Gln156		
	3	Gln215	3	Gln215
	2	Gln216	1	Gln216
	1	Gln219		
	1	Gln225		
	1	Gln303		
	1	Gln305		
	2	Lys307	1	Lys307
	3	Gln308	2	Lys307
	8	Lys315	8	Lys315
SPR1	1	Gln3		
	1	Gln4	1	Gln4
	1	Gln5		
	2	Lys6	1	Lys6
	1	Gln7	1	Gln7
	1	Gln8		
	1	Gln19		
	1	Gln82		
	1	Gln83		
	1	Lys84	1	Lys84
	2	Lys86	1	Lys86
	2	Gln87	2	Gln87
	4	Lys88	3	Lys88
SPR2			1	Gln3
	1	Gln4	1	Gln4
	1	Gln5		
	1	Gln6	1	Gln6
	1	Lys8	1	Lys8
			1	Gln9
			1	Gln12
	1	Gln63		
	1	Lys65		
	1	Lys69	1	Lys69
	3	Lys71	2	Lys71

TABLE II

Only the End Domains of SPR1 Are Used for Cross-Linking *in Vivo*

	Foreskin epidermis			Cultured keratinocytes		
CE source	Stratum corneum	Saponified	Immature	3 day	7 day	
Sequence position[a]			Numbers			Total
Gln3		1		5	3	9
Gln4			1	8	6	15
Gln5		1		2	2	5
Lys6	1		1	9	5	16
Gln7			1	4	4	9
Gln14					1	1
Gln16					1	1
Gln17			1			1
Gln18				2	1	3
Gln19		1	1	3	3	8
Lys21				1		1
Gln22				1	1	2
Gln25					1	1
Gln82				3	3	6
Gln83			1	4	2	7
Lys84			1	4	5	10
Lys86		1	1	8	5	15
Gln87	6	1	2	2	4	14
Lys88	5	2	2	4	5	18
Total	12	7	12	60	51	142

[a]These are the positions of Gln and Lys residues in the amino- and carboxy-terminal sequences only: residues in the central repeating domain were not found to be used for cross-linking *in vivo*.

crosslinks adjoining three or more sequences; 120 of 142 SPR1 and 68 of 94 SPR2 occurrences involved multiple partners (Table IV). A few examples derived from the CEs of foreskin epidermal tissue involved the SPRs and loricrin. However, most examples were recovered from the CEs of the cultured keratinocytes and involved linkages between involucrin, envoplakin, or desmoplakin. Several cases involved linkages between SPRs themselves and were further connected to other proteins.

TABLE III

Only the End Domains of SPR2 Are Used for Cross-Linking *in Vivo*

	Foreskin epidermis			Cultured keratinocytes		
CE source	Stratum corneum	Saponified	Immature	3 day	7 day	
Sequence position[a]			Numbers			Total
Gln3				5	3	8
Gln4			1	4	3	8
Gln5			1	2		3
Gln6		1	1	2	2	6
Lys8		1	1	4	3	9
Gln9		1		2	2	5
Gln12				3		3
Gln63				5	3	8
Lys65				7	6	13
Lys69		2	2	6	5	15
Lys71	8	2	4		2	16
Total	8	7	10	40	29	94

[a]These are the positions of Gln and Lys residues in the head and tail domain sequences only: residues in the central repeating domain were not found to be used for cross-linking *in vivo*.

TABLE IV

Examples of Where SPRs Function as Cross-Bridging Proteins

Item	Proteins	%SPRs: Total	Number of occurrences in CEs from[a]				
			Foreskin			Cultured cells	
			s.c. 4.5	Sapon. 0.7	Immat. 4.8	3 day 29	7 day 21
1	Loricrin-SPR1-loricrin	2			1		1
2	Loricrin-SPR2-loricrin	2			1		1
3	SPR1-loricrin-SPR1-loricrin	1		1			1
4	SPR1-loricrin-SPR2-loricrin	1			1		
5	SPR1-loricrin-SPR1-loricrin-SPR2	1			1		
6	Loricrin-SPR1-loricrin-loricrin	1		1			
7	SPR1-loricrin-SPR1-loricrin-loricrin	1		1			
8	Loricrin-SPR1-envoplakin	1				1	
9	Loricrin-SPR2-envoplakin	2				2	
10	Loricrin-SPR1-envoplakin-desmoplakin	1				1	
11	Loricrin-SPR1-involucrin-envoplakin	1				1	
12	Loricrin-SPR1-involucrin	1			1		
13	Loricrin-SPR2-involucrin	1			1		
14	Loricrin-SPR1-keratin1	2			1		1
15	Loricrin-SPR1-keratin5	3		1			1
16	Loricrin-SPR2-keratin5	2					1
17	Involucrin-SPR1-involucrin	3		1		1	1
18	Involucrin-SPR2-involucrin	4		1		2	1
19	SPR1-Involucrin-SPR1-involucrin	1				1	
20	SPR2-Involucrin-SPR1-involucrin	2				1	1
21	Involucrin-SPR1-involucrin-SPR2-involucrin	2				2	
22	Involucrin-SPR2-involucrin-SPR2-involucrin	3				2	1
23	Involucrin-SPR1-envoplakin	2				2	
24	Involucrin-SPR2-envoplakin	2				1	1
25	Involucrin-SPR2-envoplakin-desmoplakin	1					1
26	Involucrin-SPR1-desmoplakin	3			1	1	1
27	Involucrin-SPR2-desmoplakin	2				1	1
28	Involucrin-SPR1-keratin5	1				1	
29	Involucrin-SPR2-keratin1	1			1		
30	Involucrin-SPR2-keratin5	2				1	1
31	Desmoplakin-SPR1-keratin1	1				1	
32	Desmoplakin-SPR1-keratin5	2				1	1
33	Desmoplakin-SPR1-envoplakin-keratin1	1				1	
34	Desmoplakin-SPR1-envoplakin-keratin5	2				1	1
35	SPR1-SPR1-involucrin	3				2	1
36	SPR1-SPR2-involucrin	1				1	
37	SPR1-SPR1-envoplakin	1				1	
38	SPR2-SPR1-loricrin	1					1
39	SPR1-SPR1-involucrin-keratin5	1				1	
40	SPR1-SPR1-envoplakin-desmoplakin	1				1	
41	SPR1-SPR1-envoplakin-keratin5	1					1
42	Involucrin-SPR1-SPR1-loricrin	2			1		1
43	Involucrin-SPR1-SPR1-envoplakin	1				1	
44	Involucrin-SPR1-SPR1-involucrin	3				1	2
45	Involucrin-SPR1-SPR2-involucrin	2				1	1
46	Involucrin-SPR2-SPR2-involucrin	2				1	1
47	Loricrin-SPR1-SPR2-envoplakin	1				1	
48	Desmoplakin-SPR1-SPR1-involucrin	1				1	
49	Desmoplakin-SPR1-SPR2-involucrin	1					1
50	Desmoplakin-SPR1-SPR2-keratin5	1				1	

[a]Date are from s.c., stratum corneum (Steinert and Marekov, 1995); sapon., saponified foreskin stratum corneum; immat., immature foreskin epidermis (Steinert and Marekov, 1997); 3 day, CEs from 3-day cultures; 7 day, CEs from 7-day cultures of human epidermal keratinocytes (PMS and LNM, unpublished data). Note that 48 other peptides were recovered in CE digests in which an SPR was crosslinked to only one other protein partner, so that a cross-bridging role for the SPR was not evident.

Moreover, examination of the data revealed that in every case involving three or more crosslinked sequence partners, the SPR formed a cross-bridge between them. Examples of this cross-bridging role are presented in Table V. In several examples listed, nearby Gln and Lys residues on both the amino- or carboxy-termini were used multiple times. In each example, we cannot determine whether the adjacent sequences were contributed by the same or separate protein molecules, that is, whether the crosslinks were interchain involving separate SPR molecules or intrachain involving the same SPR molecule, or both. Specifically, in the case of example 3, it may be possible that residues 3–9 and residues 16–24 were contributed by the same SPR1 molecule, and in turn cross-bridged with the same or different loricrin molecules; likewise, in example 7, residues Gln4, Lys6, and Lys88 may have been contributed by the same SPR1 molecule. The reason for this uncertainty is that the crosslinked peptides were generated by extensive proteolysis of the CEs, so that only short fragments adjoined by one or more crosslinks of their full-length sequences were recovered. However, in the case of example 46, it is clear that three adjacent residues on the carboxy-terminus of an SPR2 molecule formed interchain cross-bridges with a distinctly separate SPR2 molecule, as well as with involucrin. We have not identified an example in which we can unequivocally demonstrate that intrachain crosslinking of an SPR molecule had occurred, but this possibility cannot be excluded. In many other examples, the SPR protein appeared to serve as an interchain cross-bridge between itself and its partner (Table V, examples 17, 33, and 38): again these simple cases may have arisen as a result of extensive proteolysis. In addition, on several occasions we noted that two identical SPR sequences were involved in crosslinks (examples 5, 42), so that separate SPR molecules must have participated. Together, these data indicate that multiple adjacent Gln and Lys residues on the amino- or carboxy-termini, or both, of SPR molecules can be involved in crosslinking simultaneously in a crossbridging mode with other CE proteins, including themselves.

DISCUSSION

CE structures serve vital roles in barrier function in all stratified squamous epithelia. To date, most studies have involved the epidermal CE since mutations in the genes encoding various structural proteins are prime candidates for genetic diseases of cornification with a likely ichthyosis phenotype (Williams and Elias, 1987). Since many of the same structural proteins and crosslinking TGases are used in the construction of CE structures in internal epithelia, concepts learned from epidermis may be applicable to other tissues or experimental systems such as cultured keratinocytes, and *vice versa*.

Current models have proposed two principal stages of assembly of the protein portion of the epidermal CE (Eckert *et al*., 1993; Reichert *et al*., 1993; Robinson *et al*., 1997; Steinert, 1995; Steinert and Marekov, 1995, 1997). The first "initiation" stage may involve the deposition of certain soluble proteins such as involucrin near the site where keratin intermediate filaments meet desmosomes, to form a "scaffold." This idea is supported in part by the fact that involucrin seems to be a ubiquitous component of many if not all CEs. It is proposed that this scaffold may then serve as a template for the second "reinforcement" stage of CE assembly, a process that seems to vary widely between epithelial cell types. For example, loricrin is by far the most abundant reinforcement protein in epidermal CEs, yet it is poorly expressed in cultured keratinocytes, and it is absent in many internal epithelia (Hohl *et al*., 1991, 1993). Similarly, the amounts of the SPR proteins vary widely, from essentially absent in newborn and adult interfollicular human or rodent epidermis, to modest amounts in fetal periderm (Kartasova *et al*., 1996) and neonatal human foreskin epidermis (Steinert and Marekov, 1997), to abundant amounts in mouse epidermis of the lip, snout, footpad, hair follicle, and rodent forestomach epithelium (Jarnik *et al*., 1996; Kartasova *et al*., 1996). In addition, in cultured keratinocytes and many internal epithelia such as the oral epithelia, esophagus, vagina, etc., the SPRs are the major CE components. Thus various epithelia select different structural proteins for the reinforcement stage, among which commonly are variable amounts of SPRs.

Our method of analysis of CE structure has been to use controlled proteolysis to recover peptides that contain sequences from identifiable protein components that were crosslinked together by isopeptide bonds. This method is required because it is not possible to specifically cleave the isopeptide bond alone to release the intact proteins that were associated together. Thus fragments of sequences have been obtained which we have used to reconstruct the way in which the SPR proteins were crosslinked.

Our present analyses provide novel insights into the functions of the SPRs in CE structures, summarized schematically in Fig. 1. First, we noted that only those Gln and Lys residues located on the amino- or carboxy-termini were used for crosslinking. On the other hand, the many residues within the central peptide repeats were not used, irrespective of the abundance of the SPRs in the CE samples explored. This may reflect accessibility for the TGases, or may indicate that the central repeating motifs have a separate function. Second, we noted

TABLE V
SPRs Function as Cross-Bridging Proteins in CE

Table IV item number	Sequence	Protein
Cross-bridging potentially involving the same SPR molecule		
3.	Q Q Q K Q P C ⋮ S G Q A V Q C Q S Y ⋮ Q Q Q Q V K Q P C ⋮ P T W P S K	SPR1 Lys6 Loricrin Gln153/Gln156 SPR1 Gln19/Lys21 Loricrin Lys315
7.	Q Q K T K Q K ⋮ S S Q Q V T Q T S C A ⋮ S S Q Q Q K Q P C ⋮ H Q T Q Q K Q A P T W ⋮ P T W P S K	SPR1 Lys88 Loricrin Gln215/Gln219 SPR1 Gln4/Lys6 Loricrin Gln307/Lys308 Loricrin Lys315
22.	L G L P E Q Q V L ⋮ Q Q Q Q C K Q P C ⋮ L E Q E E K L E L ⋮ P P K S K	Involucrin Gln369 SPR2 Gln6/Lys8 Involucrin Gln465/Lys468 SPR2 Lys71
37.	Q Q Q Q V K Q P C ⋮ S S Q Q Q K Q P C ⋮ A Q Q L L Q D E S S Y	SPR1 Lys21 SPR1 Gln4/Lys6 Envoplakin Gln1970
44.	E V P E Q Q V Q P K N L ⋮ Q Q K T K Q K ⋮ S S Q Q Q K Q P C ⋮ G L P E Q Q V L	Involucrin Gln455 SPR1 Lys86/Lys88 SPR1 Gln5/Lys6 Involucrin Gln368
Cross-bridging involving the same SPR molecule		
46.	P E L P E Q Q M ⋮ P S P P C Q P K C P P K S K ⋮ ⋮ P P K S K ⋮ ⋮ E H L E Q Q V G Q L	Involucrin Gln308 SPR2 Gln63/Lys65/Lys71 SPR2 Lys71 Involucrin Gln425
Simple cross-bridging		
17.	E H L Q Q V ⋮ S S Q Q Q K Q P C ⋮ P T W P S K	Involucrin Gln425 SPR1 Gln5/Lys6 Loricrin Lys315
33.	I L T C P K T K ⋮ Q Q K T K Q K ⋮ A Q L L Q D E S S Y ⋮ G G S K S I S I	Desmoplakin Lys1659 SPR1 Lys84/Gln87 Envoplakin Gln 1970/Gln1973 Keratin 1 Lys73
38.	Q Q Q Q C K Q P C ⋮ Q Q K T K Q K ⋮ P T W P S K	SPR2 Lys8 SPR1 Gln83/Gln87 Loricrin Lys315

TABLE V—*Continued*

Table IV item number	Sequence	Protein
Cross-bridging by separate SPR molecules		
5.	Q Q K T K Q K ⋮ S S Q Q V T Q T S C A ⋮ Q Q K T K Q K ⋮ H Q T Q Q K Q A ⋮	SPR1 Lys88 Loricrin Gln215/Gln219 SPR1 Lys86/Lys88 Loricrin Gln305/Gln308 SPR2 Lys71
42.	P P K S K G L P E Q Q V ⋮ Q Q K T K Q K ⋮ Q Q K T K Q K ⋮ S S Q Q V	Involucrin Gln368 SPR1 Lys86/Gln87 SPR1 Lys86/Lys88 Loricrin Gln215

that Gln and Lys residues of both the amino- and carboxy-termini were approximately equally employed in crosslinking (Tables II and III), especially for their most frequent crosslinking partners loricrin and involucrin. This means that termini of SPRs are equivalent in crosslinking, and that the SPRs molecules are functionally nonpolar. Furthermore, often several adjacent or nearby residues on the terminal sequence of the same SPR molecule were used simultaneously (Table V). Third, we have provided many examples which demonstrate that the SPRs function primarily as cross-bridging proteins between a variety of other CE structural proteins (Table IV). In CEs where the SPRs are the major protein component, as in those of cultured keratinocytes, they were extensively cross-bridged between themselves. In other CE structures where they are lesser components, they formed extensive cross-bridges between other CE proteins. Generally these were interchain cross-bridges. However, given the multiplicity of crosslinking sites on both termini, we cannot exclude the possibility of intrachain crosslinking between sequences on their amino- and/or carboxy-termini, while as well cross-bridging other CE proteins (Fig. 1). These data confirm and extend earlier notions on the roles of SPRs (Greco *et al*., 1995; Jarnik *et al*., 1996; Robinson *et al*., 1997; Steinert and Marekov, 1995). However, our data for the first time document the complexity and promiscuity of their participation in cross-bridging by use of a multiplicity of Lys and Gln residues on their terminal domains only.

Moreover, our data of Table V indicate that SPRs are utilized in three related ways in CEs of stratified squamous epithelia. First, SPRs serve in an important way to interconnect the scaffold proteins such as involucrin, cystatin α, envoplakin, and desmoplakin. As these proteins appear to be common components of the CEs of many stratified squamous epithelia, this role of SPRs likewise may be ubiquitous. Second, they function as supramolecular cross-bridging proteins by spanning between these scaffold CE proteins, and the major reinforcement CE protein loricrin in the highly specialized case of the epidermis, and furthermore, function to cross-bridge between loricrin. Third, based on our data with the CEs from the 3-day cultured keratinocytes in particular, we propose that in those epithelial tissues which do not use loricrin as the CE reinforcement protein, the highly cross-bridged network of the SPRs itself serves as the reinforcement system for the respective CE.

We note that the various SPR members differ from one another in size by variable numbers of proline-rich peptide repeats of their central domains— from as few as about 3 in the smallest SPR2 molecule, to 10–15 for the SPR1 and larger SPR2 molecules, to about 23 for the SPR3 molecule (Gibbs *et al*., 1993; Kartasova *et al*., 1988, 1996; Marvin *et al*., 1992). Based on these data and the present observations on the behavior of SPRs, we can conclude that the different SPRs may span different lengths in their cross-bridging roles between the various CE protein partners. In addition, based on commercially available secondary structure prediction algorithms, we have speculated that the three types of proline-rich peptide repeats present in the three classes of SPR proteins may have different degrees of flexibility (Kartasova *et al*. 1995). Together, therefore, we suggest that the SPRs serve as highly versatile multi-

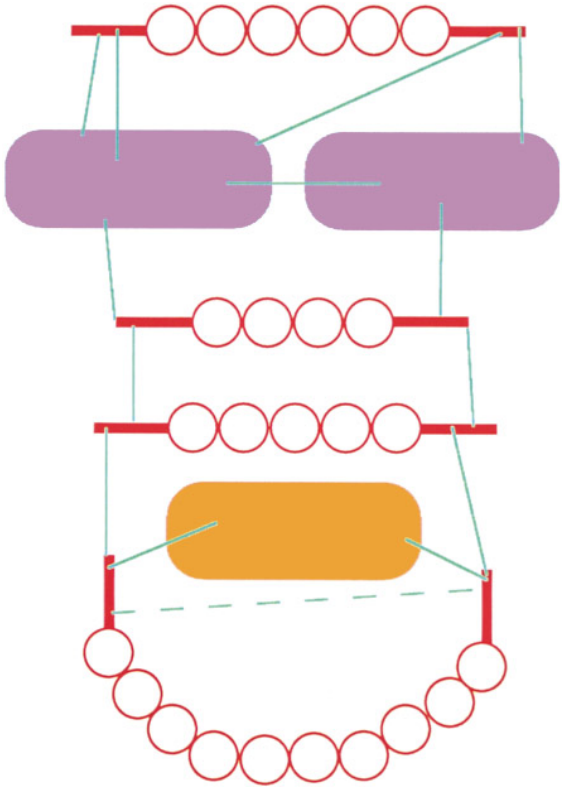

FIG. 1. Schematic drawing of the multifunctional cross-bridging roles of SPRs. A generic SPR molecule (red) contains amino- and carboxy-termini, through which all isopeptide crosslinking (shown as green rods) occurs, as well as a central domain consisting of variable numbers of peptide repeats of unknown configuration. This drawing reconstructs the patterns of crosslinks recovered for SPRs in several types of CE samples, since, typically, peptides of only 5–15 amino acids of the intact proteins were recovered. Our data reveal that the amino- and carboxy-termini are functionally equivalent; that is, SPR cross-bridging occurs in a nonpolar orientation. Our data indicate that several Gln or Lys residues may be used on either terminus simultaneously. The SPRs form interchain cross-bridges between different CE molecules (pink or orange ovoids) and/or between themselves. Although we have not recovered a single example, there is no *a priori* reason why a single SPR molecule could not form an intrachain cross-bridge within either one of its termini or involving both termini (dashed green line). It must be pointed out that the CE structure is three-dimensional, packed at high density, and extensively crosslinked, involving multiple "scaffold" or "reinforcement" proteins. Moreover, the tertiary structure of none of these proteins is known at this time.

functional cross-bridging proteins within CEs. Their ubiquitous and highly variable expression in different epithelia indicates that they may play important roles in defining CE structure in tissue-specific ways.

REFERENCES

An, G., Tesfaigzi, J., Chuu, Y. J., and Wu, R. (1993) Isolation and characterization of the human spr1 gene and its regulation of expression by phorbol ester and cyclic AMP, *J. Biol. Chem.* **268,** 10977–10982.

Austin, S. J., Fujimoto, W., Marvin, K. W., Vollberg, T. M., Lorand, L., and Jetten, A. M. (1996) Cloning and regulation of cornifin beta, a new member of the cornifin/spr family. Suppression by retinoic acid receptor-selective retinoids, *J. Biol. Chem.* **271,** 3737–3742.

Candi, E., Tarcsa, E., DiGiovanna, J. J., Compton, J. G., Elias, P. M., Marekov, L. N., and Steinert, P. M. (1998) Identification of a highly conserved lysine residue on type II keratins which is essential for the structural coordination of keratin intermediate filaments and the cornified cell envelope through isopeptide crosslinking by transglutaminases. *Proc. Natl. Acad. Sci. USA*, in press.

Candi, E., Melino, G., Pei, G., Tarcsa, E., Marekov, L. N., and Steinert, P. M. (1995) Bacterially-expressed human loricrin: Biochemical, structural and transglutaminase substrate properties of the major epidermal cornified cell envelope structural protein, *J. Biol. Chem.* **270,** 26382–26390.

Downing, D. T., Stewart, M. E., Wertz, P. W., and Strauss, J. S. (1993) Lipids in the epidermis and sebaceous glands, *in* Fitzpatrick, T. B., Eisen, A. Z., Wolff, K., Freedberg, I. M., and Austen, K. F. (Eds.), Dermatology in General Medicine, pp. 210–221, McGraw-Hill, New York.

Eckert, R. L., Yaffe, M. B., Crish, J. F., Murthy, S., Rorke E. A., and Welter, J. F. (1993) Involucrin-structure and role in envelope assembly, *J. Invest. Dermatol.* **100,** 613–617.

Elias, P. M. (1996) The stratum corneum revisited, *J. Dermatol.* **23,** 756–758

Fujimoto, W., Nakanishi, G., Arata, J., and Jetten, A. M. (1997) Differential expression of human cornifin α and β in squamous differentiating epithelial tissues and several skin lesions, *J. Invest. Dermatol.* **108,** 200–204.

Gibbs, S., Fijneman, R., Wiegant, J., van Kessel, A. G., van de Putte, P., and Backendorf, C. (1993) Molecular characterization and evolution of the SPRR family of keratinocyte differentiation markers encoding small proline rich proteins, *Genomics* **16,** 630–637.

Greco, M. A., Lorand, L., Lane W. S., Baden, H. P., Parameswaran, K. N. P., and Kvedar, J. C. (1995) The pancornulins: A group of small proline-rich-related cornified envelope precursors with bifunctional crosslinking capabilities and isopeptide bond formation, *J. Invest. Dermatol.* **104,** 204–210.

Greenberg, C. S., Birckbichler, P. J., and Rice, R. H. (1991) Transglutaminases: Multifunctional crosslinking enzymes that stabilize tissues, *FASEB J.* **5,** 3071–3077.

Hohl, D. (1990) Cornified cell envelope, *Dermatologica* **180,** 201–211.

Hohl, D., Lichti, U., Turner, M. L., Roop, D. R., and Steinert, P. M. (1991) Characterization of human loricrin: Structure and function of a new class of epidermal cell envelope proteins, *J. Biol. Chem.* **266,** 6626–6636.

Hohl, D., Ruf, O. B., de Viragh, P. A., Huber, M., Detrisac, C. J., Schnyder, U. W., and Roop, D. R. (1993) Expression patterns of loricrin in various species and tissues, *Differentiation* **54,** 25–34.

Hohl, D., de Viragh, P. A., Amiguet-Barras, F., Gibbs, S., Backendorf, C., and Huber, M. (1995) The small proline rich proteins constitute a multigene family of differentially regulated cornified cell envelope precursor proteins, *J. Invest. Dermatol.* **104,** 902–909.

Huber, M., Rettler, I., Bernasconi, K., Frenk, E., Lavrijsen, S. P. MS. P., Ponec, M., Bon, A., Lautenschlager, S., Schoderer, D. F., and Hohl, D. (1995) Mutations in keratinocyte transglutaminase in lamellar ichthyosis, *Science* **267**, 525–528.

Jarnik, M., Kartasova, T., Steinert, P. M., Lichti, U., and Steven, A. C. (1996) Differential expression and cell envelope incorpora-

tion of small proline rich protein 1 in different cornified epithelia, *J. Cell Sci.* **109,** 1381–1391.

Kartasova, T., van Muijen G., Pelt-Heerschap, H., and van de Putte, P. (1988) Isolation, characterization, and UV-stimulated expression of two families of genes encoding polypeptides of related structure in human epidermal keratinocytes, *Mol. Cell Biol.* **8,** 2195–2210.

Kartasova, T., Parry, D. A. DD. A., and Steinert, P. M. (1995) Modeling of small proline rich protein evolution and structure, *J. Invest. Dermatol.* **104,** 611.

Kartasova, T., Kohno, Y., Koizumi H., Osada, S., Huh, N., Lichti, U., Steinert, P. M., and Kuroki, T. (1996) Sequence and expression patterns of mouse SPR1: Correlation of expression with epithelial function, *J. Invest. Dermatol.* **106,** 294–304.

Martinet, N., Beninati, S., Nigra, T. P., and Folk, J. E. (1988) N^1N^8-Bis(γ- glutamyl)spermidine cross-linking in epidermal cell envelopes, *Biochem. J.* **271,** 305–308.

Marvin, K. W., George, M. D., Fujimoto, W., Saunders, N. A., Bernacki, S. H., and Jetten, A. M. (1992) Cornifin, a cross-linked envelope precursor in keratinocytes that is down-regulated by retinoids, *Proc. Natl. Acad. Sci. USA* **89,** 11026–11030.

Owens, D. M., Zainal, T. A., Jetten, A. M., and Smart, R. C. (1996) Localization and expression of cornifin alpha/SPRR1 in mouse epidermis, anagen hair follicles, and skin neoplasms, *J. Invest. Dermatol.* **106,** 647–654.

Reichert, U., Michel, S., and Schmidt, R (1993) The cornified cell envelope: A key structure of terminally differentiating keratinocytes, *in* Darmon, M., and Blumenberg, M. (Eds.), Molecular Biology of the Skin, pp. 107–150, Academic Press, New York.

Robinson, N. A., Lapec, S., Welter, J. F., and Eckert, R. L. (1997) S100A11, S100A10, annexin I, desmosomal proteins, small proline-rich proteins, plasminogen activator inhibitor-2, and involucrin are components of the cornified envelope of cultured human epidermal keratinocytes, *J. Biol. Chem.* **272,** 12035–12046.

Russell, L. J., DiGiovanna, J. J., Hashem, N., Rogers, G. R., Steinert, P. M., Compton, J. G., and Bale, S. J. (1995). *TGM1* mutations in lamellar ichthyosis, *Nature Genet.* **9,** 279–283.

Simon, M. (1994) The epidermal cornified cell envelope and its precursors, *in* Leigh, I. M., Lane E., and Watt, F. M. (Eds.), The Keratinocyte Handbook, pp. 275–292, Cambridge Univ. Press, Cambridge.

Steinert, P. M. (1995) A model for the hierarchical structure of the human epidermal cornified cell envelope, *Cell Death Different.* **2,** 23–31.

Steinert, P. M., and Marekov, L. N. (1995) The proteins elafin, filaggrin, keratin intermediate filaments, loricrin and SPRs are isodipeptide crosslinked components of the human epidermal cornified cell envelope, *J. Biol. Chem.* **270,** 17702–17711.

Steinert, P. M., and Marekov, L. N. (1997) Involucrin is an important early component in the assembly of the epidermal cornified cell envelope, *J. Biol. Chem.* **272,** 2021–2030.

Steinert, P. M., Kartasova, T., and Marekov, L. N. (1998) Submitted.

Steven, A. C., and Steinert, P. M. (1994) The protein composition of the cornified cell envelope, *J. Cell Sci.* **107,** 693–700.

Swartzendruber, D. C., Wertz, P. W., Madison, K. C., and Downing, D. T. (1987) Evidence that the corneocyte has a chemically bound lipid envelope, *J. Invest. Dermatol.* **88,** 709–713.

Tarcsa, E., Marekov, L. N., Andreoli, J. M., Idler, W. W., Candi, E., Chung, S.-I., and Steinert, P. M. (1997) The fate of trichohyalin: Sequential post-translational modifications by peptidylarginine deiminase and transglutaminases, *J. Biol. Chem.* **272,** 27893–27901.

Williams M. L., and Elias, P. M. (1987) Genetically transmitted, generalized disorders of cornification: the ichthyoses, *in* Alper, J. C. (Ed.), Dermatology Clinics, Vol. 5, pp. 155–178, Saunders, Philadelphia.

Yoneda, K., and Steinert, P. M. (1993) The over-expression of loricrin in transgenic mice produces a normal phenotype. *Proc. Natl. Acad. Sci. USA* **90,** 10754–10758.