Supp. Table S1. Isoforms of collagen IV

Chain	Isoform	Residues	Canonical sequence	Replacement
α1(IV)	Isoform 2	498 - 848		Missing
α3(IV)	Isoform 2	1586 - 1670	FTSAG... MKKRH	KAYSINCESWGIRKNNKSLSG VHEEKTLKLKKTAELVFFILKNK VMTEHAVI
	Isoform 3	1488-1670	GTLGS...M KKRH	DALFVKVLRSP
	Isoform 4	1488-1670	GTLGS...M KKRH	ESLFHQL
	Isoform 5	1418 - 1424	GPAGSDG	ESLFFHQL
		1425 - 1670		Missing
α5(IV)	Isoform 2	1264-1264	G	GPTGFQG, longer isoform found in the kidney
α6(IV)	Isoform 2	5	MLINK	Missing

Supp. Table S2. Predicted* integrin-binding sites in collagen IV

	Residues					
Motif	**α1 chain**	**α2 chain**	**α3 chain**	**α4 chain**	**α5 chain**	**α6 chain**
GFPGER	385 - 390		756 - 761	915 – 920 984 - 989	400 - 405	
GAPGER		1257 - 1262				
GFPGEK						596-600
GLPGEK	676 – 681 900 – 905 1112 – 1117	1347 - 1352			365 - 370	605-610

* These have generally not been confirmed

Supp. Table S3. Location of binding sites for laminin and other extracellular structural molecules

Proposed laminin binding sites – location or distance from NCI						**Reference**
					NC1	Rao et al., 1985
7S domain				*86 nm		Charonis et al., 1985
	251 nm		174 nm	87 nm		Ohno et al., 1991
	291 nm	216 nm		81 nm		Laurie et al., 1986
			178 nm	75 nm		Aumailley et al., 1989
Summary of proposed laminin binding sites						
7S domain	251-291nm (303 aa)	216 nm (534 aa)	174 – 178 nm (702 aa)	75 – 87 nm (1100 aa)	NC1	
Overlap with other extracellular matrix binding sites – distance from NCI						
		Fibronectin (205 nm); HSPG (206 nm)		Nidogen (75 nm); HSPG (82 nm)		Other binding sites (Laurie et al., 1986; Paralkar et al., 1990)

*Distances have been corrected so they are all measured in an amino direction from the intersection of the triple helix and carboxy NC1 domain.

Supp. Table S4a. Missense mutations in the α5 chain of collagen IV associated with early ('severe') or late onset renal failure in X-linked Alport syndrome – effect of substituting residue

Gly substitutions	Number of missense mutations identified (%)	Number of mutations with severe disease (%)	Number of mutations with mild disease (%)	Number with severe: mild disease
Arg - R	44 (25%)	16 (36%)	5 (11%)	3.3
Cys - C	9 (5%)	2 (22%)	3 (33%)	0.7
Ser - S	23 (13%)	6 (26%)	8 (35%)	0.7
Ala -A	7 (4%)	2 (29%)	1 (14%)	2.1
Asp -D	29 (16%)	7 (24%)	5 (17%)	1.4
Glu - E	32 (18%)	14 (44%)	3 (9%)	4.9
Val – V	34 (19%)	13 (38%)	3 (9%)	4.2
Trp - W	1 (1%)	0 (0%)	1 (100%)	N/A
Pro - P	0 (0%)	0 (0%)	0 (0%)	N/A
Total	179	60 (34%)	29 (16%)	1.2

Supp. Table S4b. Missense mutations in the α5 chain of collagen IV associated with early or late ('severe') onset renal failure in X-linked Alport syndrome – effect of location and ligand-binding site

Disease-associated variants in α5 chain	Total number	Number with severe disease	Number without severe disease or severity not specified	Severe/mild
Total	209	73	32	2.3
Amino one third (to residue 550)	27	11	4	2.8
Collagenous domain	179	60	29	2.1
Noncollagenous residues	30	13	3	4.3
NC1 domain	20	6	3	2.0
Exons 25 and 26		15	6	2.5
Integrin-binding sites	2	2	None	
Heparin-binding sites	3	1	2	0.3

Clinical phenotype (severe = early onset renal failure) was not available for all missense mutations in X-linked Alport syndrome and this table indicates whether mutations affecting a certain location or ligand-binding site were more likely to result in severe disease.