

Supplementary materials:

Table 1: Functional properties of human collagens in diseases:

Collagen	Accession number	Collagen name	Function	Involvement in disease
Col 1	P02452	CO1A1_HUMAN Collagen alpha-1(I) chain	Fibrillar forming collagen in bones tendons ligaments	Mutation results in Dwarfism,Ehlers-Danlos syndrome and Osteogenesis imperfecta
Col 2	P02458	CO2A1_HUMAN Collagen alpha-1(II) chain	Normal embryonic development of skeleton,linear growth	Cataract,deafness and mutation causes stickler syndrome and dwarfism
Col 3	P02461	CO3A1_HUMAN Collagen alpha-1(III) chain	Fibrillar forming collagen like type 1 collagen,formation of connective tissue	Mutation causes Aortic aneurysm, Ehlers-Danlos syndrome
Col 4	P02462	CO4A1_HUMAN Collagen alpha-1(IV) chain	Formation of 'chicken wire' meshwork in glomerulus tissue ,thus involved in filtering of urine in kidney	Alport syndrome
Col 5	P20908	CO5A1_HUMAN Collagen alpha-1(V) chain	Fibrillar forming,binds to dna,fibrin,heparin and insulin	Ehlers-Danlos syndrome
Col 6	P12109	CO6A1_HUMAN Collagen alpha-1(VI) chain	Cell binding protein	Ulrich myopathy and Bethlem myopathy
Col 7	Q02388	CO7A1_HUMAN Collagen alpha-1(VII) chain	Basement membrane organization and adherence	epidermolysis bullosa dystrophica
Col 8	P27658	CO8A1_HUMAN Collagen alpha-1(VIII) chain	Proliferation of vscular smooth muscle cells,maintains vessel wall integrity and structure	Posterior polymorphous corneal dystrophy 2
Col 9	P20849	CO9A1_HUMAN Collagen alpha-1(IX) chain	Flexible hence connects type II collagen to other cartilage components	Osteoarthritis,muation may result in epiphyseal dysplasia and stickler syndrome
Col 10	Q03692	COAA1_HUMAN Collagen alpha-1(X) chain	-	Schmid type metaphyseal chondrodysplasia
Col 11	P12107	COBA1_HUMAN Collagen alpha-1(XI) chain	Important role in fibrillogenesis by controlling growth of collagen II fibrils	Marshall syndrome and stickler syndrome
Col 12	Q99715	COCA1_HUMAN Collagen alpha-1(XII) chain	Modifies surrounding matrix by interaction with type I collagen	N.D
Col 13	Q5TAT6	CODA1_HUMAN Collagen alpha-1(XIII) chain	Cell-matrix and cell-cell interactions for normal development	-
Col 14	Q05707	COEA1_HUMAN Collagen alpha-1(XIV) chain	Interacts with collagen bundles,adhesion	-
Col 15	P39059	COFA1_HUMAN Collagen alpha-1(XV) chain	Stabilizes microvessels and muscle cells in heart and skeletal muscle,inhibits angiogenesis	-
Col 16	Q07092	COGA1_HUMAN Collagen alpha-1(XVI) chain	Induces integrin mediated interactions:Cell spreading, attachment and alteration of morphlogy of cells	-
Col 17	Q9UMD9	COHA1_HUMAN Collagen alpha-1(XVII)	Role in maintaining integrity of hemidesmosome, attachment of basal keratinocytes basement membrane.	-
Col 18	P39060	COIA1_HUMAN Collagen alpha-1(XVIII) chain	Determination of renal structure and closure of renal tube,inhibits angiogenesis	-

Col 19	Q14993	COJA1_HUMAN Collagen alpha-1(XIX) chain	Esophagus development,organization of pericellular matrix	-
Col 20	Q9P218	COKA1_HUMAN Collagen alpha-1(XX) chain	-	-
Col 21	Q96P44	COLA1_HUMAN Collagen alpha-1(XXI) chain	-	-
Col 22	Q8NFW1	COMA1_HUMAN Collagen alpha-1(XXII) chain	Acts as a cell adhesion ligand for skin epithelial cells and fibroblasts	-
Col 23	Q86Y22	CONA1_HUMAN Collagen alpha-1(XXIII) chain	-	-
Col 24	Q17RW2	COOA1_HUMAN Collagen alpha-1(XXIV) chain	Regulation of type I collagen fibrillogenesis at specific anatomical locations during fetal development	-
Col 25	Q9BXS0	COPA1_HUMAN Collagen alpha-1(XXV) chain	Binds heparin,assembles amyloid fibrils into protease resisitant aggregates	-
Col 26	Q96A83	EMID2_HUMAN Collagen alpha-1(XXVI) chain	-	-
Col 27	Q8IZC6	CORA1_HUMAN Collagen alpha- 1(XXVII) chain	calcification of cartilage and the transition of cartilage to bone	-
Col 28	Q2UY09	COSA1_HUMAN Collagen alpha- 1(XXVIII) chain	May act as a cell binding protein	-

*N.D. - Not defined

Table 2: Amino acid composition of human alpha-1 collagens (in %) :

Collagen	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
Col 1	9.5	4.8	1.2	4.5	1.2	3.3	5.1	26.7	0.6	1.6	3.3	3.9	0.9	1.8	19	4.1	3.1	0.4	0.9	3.2
Col 2	9	4.8	2.2	4.2	1.3	4	5.3	27.3	0.5	2.3	3.8	4.5	1.1	1.7	18.2	3.2	3.0	0.5	0.7	2.6
Col 3	7.8	4.1	2.8	3.8	1.5	2.9	5.0	28.2	1.0	2.5	3.3	4.2	1.2	1.6	19.2	5.0	2.1	0.5	0.1	2.5
Col 4	3.5	2.7	1.0	3.5	1.2	4.4	4.2	28.6	1.0	3.5	5.5	5.6	1.9	2.8	19.4	4.3	2.6	0.4	1.1	3.1
Col 5	5	3.8	1.8	5.7	0.7	4.0	6.5	23.3	0.9	2.8	5.3	5.3	1.3	2.1	18.2	3.8	3.8	0.4	2.2	2.9
Col 6	7.5	5.8	2.7	6.9	1.9	4.1	6.6	15.2	1.4	4.1	6.9	5.3	1.1	3.2	8.7	5.3	4	0.4	2.7	6.3
Col 7	6.6	7.4	0.9	4.9	0.6	3.6	6.4	21.3	1.0	1.9	7.2	3.2	0.6	1.3	14.5	5.6	4.5	0.6	1.2	6.8
Col 8	4.4	1.7	1.1	1.5	0.3	5.1	3.5	25.9	1.5	4.4	7.1	6.3	3.1	2.3	22.4	1.5	1.1	0.1	3.0	4.0
Col 9	5.1	5.9	2.2	4.2	1.2	4.2	5.1	23.6	0.9	3.9	6.2	4.6	1.3	2.4	16.6	4.6	2.8	0.7	0.7	4.0
Col 10	5.3	2.8	2.1	1.8	0.1	3.4	3.2	25.7	1.5	4	5.1	5.1	1.6	2.2	21.3	4	3.5	0.3	3.1	3.8
Col 11	5.4	3.8	2.0	5.5	0.6	4.5	6.8	23.4	0.9	3.1	4.4	5.9	1.3	2.6	15.9	4	3.8	0.5	2.0	3.5
Col 12	5.2	5.1	3.9	5.6	0.7	3.4	6.3	9.2	1.0	4.8	6.6	5.1	1.5	3.0	8.6	7.9	8.7	1.0	3.7	8.8
Col 13	6.8	4.9	1.5	3.2	1.1	3.9	6.1	25.7	1.5	2.6	7.4	6.4	1.8	0.7	17.4	3.3	2.4	0.3	0.4	2.4
Col 14	5.2	4.1	3.2	5.2	1.1	4.0	6.6	10.9	1.6	5.5	7.2	4.8	1.8	3.3	8.7	7.4	7.6	1.0	3.0	7.9
Col 15	7.8	3.4	2.7	4.2	0.7	2.9	7.0	15.9	1.8	3.9	8.1	3.7	2.2	3.1	13.8	7.1	5.3	0.7	1.0	4.7
Col 16	5.3	3.9	1.4	3.2	2	4.9	5.7	24.4	1.2	2.6	7.1	5.4	1.6	1.9	17.6	4.7	2.7	0.5	0.8	3.9
Col 17	5.7	4.5	2.1	3.6	0.5	3.2	4.2	18.8	1.5	2.7	8.4	4	2.2	1.7	13.6	11.9	5.6	0.7	2.4	3.9
Col 18	8	5.2	1.5	4.2	1.3	3.8	5.1	17.2	2.1	1.7	6.2	2.4	0.9	3.2	16.9	6.9	3.9	1.2	1.1	5.0
Col 19	4.6	4.1	2.6	4.3	1.2	4	5.9	22.9	1.4	5	10.7	6.7	1.6	2.2	15.1	4.6	2.6	0.6	1.4	2.9
Col 20	8.9	6.6	0.9	3.6	1	4.4	5.7	11.5	2.2	1.8	6.9	3	1.1	2.7	10	8.3	6.7	1.3	2.0	7.6
Col 21	4.2	4	2.7	5	1.5	5.3	5.2	18.9	1.1	5.6	6.2	7.1	1.3	3.0	11.1	5.2	3.6	0.4	2.3	5.5
Col 22	5.8	5.2	1.5	4.3	1.1	3.8	6.4	24.4	1.2	2.8	7.4	5.4	1.2	2.1	16.4	4.1	2.6	0.4	1.0	4.3
Col 23	8	3.9	0.4	5.4	1.1	3.1	6.7	27.2	0.6	1.5	7.5	6.5	0.9	0.4	17.2	3.1	1.5	0.4	0.4	2.8
Col 24	3.4	4.3	3.3	3.5	0.9	5.3	6	21.3	2.2	5.1	6.3	6.1	1.4	2.5	12.1	5.3	4.6	0.3	1.8	4.0
Col 25	5	6.3	1.4	4.3	1.1	4.9	6.9	25.5	1.5	3.1	8.8	7.6	2.6	0.8	16.5	3.1	2.4	0.2	0.6	2.0
Col 26	8.6	5.4	2.3	3.6	2.9	4.1	5.4	14.7	1.6	1.6	7.2	2.7	1.4	0.9	15.4	6.6	5.2	1.1	1.8	4.8
Col 27	5.9	3.6	1.1	3.7	0.8	5.2	3.6	21.3	1.7	2.4	5.9	5.1	2	2.6	16.3	5.7	4.7	0.5	0.9	3.8
Col 28	3.7	3.6	2.4	5.6	1.3	5.2	6.5	18.8	0.6	5.4	5.9	8.1	1.2	3.4	11.5	6.1	4	0.4	1.6	4.7

*Col- Collagens

Table 3(a): Physico-chemical parameters of human alpha-1 collagens:

Collagen	No. of amino acids	Molecular weight	pI	'-' charged residue	'+' charged residue	Extinction Coefficient	Instability index	Aliphatic index	GRAVY
Col 1	1464	138941.5	5.6	141	128	53495	30.43	37.98	-0.788
Col 2	1487	141785.3	6.58	141	139	54525	25.21	40.03	-0.803
Col 3	1466	138564.2	6.21	129	122	62225	30.18	37.31	-0.797
Col 4	1669	1606147.7	8.55	128	138	61070	32.04	47.39	-0.621
Col 5	1838	183559.8	4.94	225	168	98850	33.09	45.35	-0.873
Col 6	1028	108529.4	5.26	139	114	64970	28.52	68.7	-0.525
Col 7	2944	295219.6	5.95	332	310	159140	32.07	61.86	-0.625
Col 8	744	73364	9.62	37	60	38405	36.06	61.21	-0.434
Col 9	921	91869.2	8.94	86	96	42565	32.61	56.13	-0.658
Col 10	680	66157.9	9.68	34	54	42290	25.95	51.94	-0.556
Col 11	1806	181064.8	5.06	222	174	103765	30.81	44.91	-0.859
Col 12	3063	333146.7	5.38	366	313	334620	32.90	75.45	-0.427
Col 13	717	69949.9	9.27	67	81	15970	31.44	52.87	-0.765
Col 14	1796	193515.4	5.16	211	160	179095	37.57	77.67	-0.326
Col 15	1388	141720.1	4.9	155	95	76485	40.19	68	-0.377
Col 16	1604	157751.3	8.14	144	150	65370	35.88	50.73	-0.671
Col 17	1497	150419.3	8.89	117	128	109015	45.25	55.47	-0.573
Col 18	1754	178187.6	5.67	164	133	145185	48.57	61.72	-0.467
Col 19	1142	115220.7	8.57	116	124	63215	30.68	56.68	-0.708
Col 20	1284	135830	8.27	119	123	132990	45.18	79.61	-0.261
Col 21	957	99368.5	8.57	98	106	55655	33.28	69.14	-0.517
Col 22	1626	161145.3	6.88	174	172	57965	34	53.28	-0.715
Col 23	540	51943.9	6.88	65	65	14355	30.81	50.69	-0.829
Col 24	1714	175496.3	8.46	162	170	73075	28.32	64.21	-0.622
Col 25	654	64770.7	8.6	73	78	11835	24.85	47.19	-0.919
Col 26	441	45381.1	7.02	40	40	40170	46.77	63.11	-0.523
Col 27	1860	186892.3	9.83	136	196	81205	37.62	54.15	-0.637
Col 28	1125	116657.1	6.1	136	131	55195	24.18	61.42	-0.66

*Col- Collagens

Table 3(b): Physico-chemical properties of human alpha-1 collagens

Collagen	No. of codons	Bulkiness	Polarity	Refractivity	Recognit ion factors	Hydrophobicity	Trans-membrane tendency	% buried residues	% accessible residues	Average area buried	Average flexibility
Col1	3.889	14.3645	17.2635	14.4125	87.5	-0.2835	-0.6295	6.528	5.622	118.5835	0.4505
Col2	3.611	14.535	17.012	14.417	89.722	0.05	-0.434	6.9445	5.9665	116.9055	0.452
Col3	3.889	13.688	14.3925	14.1385	88.889	-0.222	-0.3985	7.3885	5.65	116.239	0.4505
Col4	3.722	13.466	16.9645	14.739	89.6115	0.0115	-0.564	6.561	5.9555	120.7055	0.4445
Col5	3.889	13.905	16.8185	15.437	88.5555	-0.089	-0.429	6.1335	5.9	117.989	0.452
Col6	3.5555	13.866	17.659	16.544	88.7225	-0.0225	-0.5395	5.95	5.672	116.35	0.453
Col7	3.7775	14.1405	17.502	14.787	90.611	-0.139	-0.6305	6.767	5.9445	118.028	0.4485
Col8	3.722	13.8405	14.092	15.435	88.389	-0.1835	-0.3875	6.5835	5.8725	119.7725	0.4455
Col9	3.5555	14.563	17.537	16.071	87.5555	-0.0055	-0.3405	6	5.422	122.5835	0.444
Col10	3.722	14.577	12.346	14.2365	88.5555	-0.2665	-0.3025	6.25	5.661	119.7725	0.4555
Col11	3.4445	13.735	17.296	16.398	89	-0.55	-0.6195	6.611	6.0665	122.972	0.45
Col12	3.722	14.226	16.8455	15.1145	90	-0.078	-0.5105	6.9225	5.9335	121.2225	0.454
Col13	3.833	13.779	20.1465	15.135	80.0555	-0.2335	-0.55	6.4615	5.9835	118.511	0.4495
Col14	3.5555	14.611	19.96	15.7995	89.3335	-0.283	-0.843	5.622	5.7055	122.35	0.4495
Col15	3.72205	13.5155	16.584	15.1545	87.444	-0.033	-0.513	6.861	5.4945	119.6555	0.443
Col16	3.5	14.316	17.226	15.108	88.2775	0.1775	-0.357	6.1055	5.789	120.672	0.448
Col17	3.889	12.439	17.232	15.038	92.333	-0.1165	-0.3895	7.1615	6.122	116.522	0.454
Col18	3.889	13.5065	17.062	14.587	89.444	0.3335	-0.3765	6.3385	5.5275	122.828	0.4335
Col19	3.5	14.179	17.4405	16.2065	87.4445	-0.622	-0.4485	6.311	5.7165	126.9225	0.4445
Col20	3.7775	14.161	19.7095	15.07	90.6665	-0.3335	-0.5825	6.828	5.9115	124.783	0.4425
Col21	3.5555	14.8025	16.9925	15.4085	88.0555	-0.3335	-0.3235	6.256	5.811	123.1165	0.439
Col22	3.722	14.055	17.253	15.4765	87.056	0.222	-0.32	6.4555	5.911	124.278	0.4435
Col23	3.9445	11.8145	19.881	11.6245	87.222	-0.017	-0.5535	5.63385	6	109.889	0.4535
Col24	3.722	14.5095	17.514	15.4565	89.1115	-0.2555	-0.472	7.1	5.8225	119.835	0.4395
Col25	3.7225	13.5035	17.207	13.0145	87.3885	0.011	-0.534	7.289	5.6555	113.683	0.4545
Col26	3.167	13.7255	14.878	15.4075	87.8885	-0.3055	-0.593	6.767	5.2225	116.1885	0.4455
Col27	3.611	13.9525	17.331	14.9185	90.889	-0.111	-0.3675	6.55	6.011	119.067	0.445
Col28	3.5555	14.7145	19.934	16.194	88.3335	-0.183	-0.5865	6.022	5.7165	127.9725	0.446

*Col- Collagens

Table 4: Secondary structural features of human alpha-1 collagens (in %):

Collagen	α helix	3_{10} helix	Pi helix	β bridge	Extended strand	β turn	Bend region	Random coil	Ambiguous states	Other states
Col1	4.37	0	0	0	8.13	4.3	0	83.2	0	0
Col2	6.27	0	0	0	7.2	4.57	0	81.51	0	0
Col3	4.5	0	0	0	7.44	4.71	0	83.36	0	0
Col4	4.73	0	0	0	7.61	5.69	0	81.97	0	0
Col5	7.94	0	0	0	11.53	4.9	0	75.63	0	0
Col6	24.22	0	0	0	15.86	5.25	0	54.67	0	0
Col7	9.51	0	0	0	15.9	6.15	0	68.44	0	0
Col8	5.65	0	0	0	10.89	9.41	0	74.06	0	0
Col9	9.45	0	0	0	7.17	3.37	0	80.02	0	0
Col10	3.97	0	0	0	12.35	5	0	78.68	0	0
Col11	8.36	0	0	0	12.02	5.59	0	74.03	0	0
Col12	13.12	0	0	0	25.56	4.9	0	56.42	0	0
Col13	14.09	0	0	0	4.6	4.46	0	76.85	0	0
Col14	17.76	0	0	0	24.44	5.57	0	52.23	0	0
Col15	18.44	0	0	0	16.71	6.77	0	58.07	0	0
Col16	7.79	0	0	0	10.1	5.99	0	76.12	0	0
Col17	19.91	0	0	0	11.89	6.95	0	61.26	0	0
Col18	17.56	0	0	0	12.43	6.9	0	63.11	0	0
Col19	13.57	0	0	0	9.54	6.04	0	70.84	0	0
Col20	18.15	0	0	0	21.65	5.69	0	54.52	0	0
Col21	15.26	0	0	0	12.43	4.7	0	67.61	0	0
Col22	11.38	0	0	0	10.7	7.56	0	70.36	0	0
Col23	13.7	0	0	0	0.56	1.3	0	84.44	0	0
Col24	7.41	0	0	0	9.33	2.92	0	80.34	0	0
Col25	12.69	0	0	0	4.43	3.82	0	79.05	0	0
Col26	24.94	0	0	0	11.11	5.9	0	58.05	0	0
Col27	9.46	0	0	0	11.88	5.86	0	72.8	0	0
Col28	19.29	0	0	0	10.58	4.18	0	65.96	0	0

*Col- Collagen

Table 5: Motifs in human alpha-1 collagens:

Collagen	Motif found	Motif ID	Description	Start	End	Match Status	Significance
Col1	VWFC_1	PS01208	VWFC domain signature	58	95	Strong match; not a false positive	
Col2	VWFC_1	PS01208	VWFC domain signature	52	89	Strong match; not a false positive	
Col3	VWFC_1	PS01208	VWFC domain signature	50	88	Strong match; not a false positive	
Col7	BPTI_KUNITZ_1	PS00280	Pancreatic trypsin inhibitor signature (Kunitz) family	2907	2925	Strong match; not a false positive	
Col28	BPTI_KUNITZ_1	PS00280	Pancreatic trypsin inhibitor signature (Kunitz) family	1100	1118	Strong match; not a false positive	

*VWFC: von Willebrand factor (VWF) type C repeat

**No motifs were found in remaining human alpha-1 collagen sequences

*** Col- Collagen