2007 VOLUME 63-64

PRINT ISSN 1843-5262
ELECTRONIC ISSN 1843-536X

BULLETIN

1869

USAMV

OF
UNIVERSITY OF AGRICULTURAL SCIENCES AND VETERINARY MEDICINE CLUJ-NAPOCA

ANIMAL SCIENCE AND BIOTECHNOLOGIES

Bulletin USAMV-CN, 63 - 64/2007

AMINO ACIDS SEQUENCE ANALYSIS ON COLLAGEN

Sorana D. Bolboacă [1], Jäntschi L [2]

[1]"Iuliu Hatieganu" University of Medicine and Pharmacy Cluj-Napoca, 6 Louis Pasteur, 400349 Cluj-Napoca, Romania, sbolboaca@umfcluj.ro
[2]Technical University of Cluj-Napoca, 103-105 Muncii Bvd., 400641 Cluj-Napoca, Romania.
lori@academicdirect.org

Key words: Amino Acids; Collagen Type I; Properties; Multivariate Analysis

Abstract. Staring from available information about amino acids properties and sequences on collagen type I chains, the aims of the study were to identify the principal property component and to analyze the similarities within and between collagens on five species. The principal component analysis applied on twenty-four amino acids properties revealed that the hydrophobic or hydrophilic character measured by Wealling et al. is more stable comparing with the other investigated properties. Similarity analysis identified similar and dissimilar within and between studied species from the viewpoint of amino acids sequences on collagen type I alpha 1 and 2 chains.

INTRODUCTION

Collagen, the main protein of connective tissues, is a fibrous protein, inextensible, which can be found at the level of connective tissues from heart, vessels, skin, cornea, cartilage, ligaments, tendons, bone, and teeth. Today are known twenty-eight types of collagens (Veit et al., 2006).

The structure of collagen type I, distributed at the level of all tissues in organism, is known from a long time (Kadler et al., 1996). The COL1A1 and COL1A2 genes which coding the collagens chains have also been identified and characterized: *Homo Sapiens* (Chu et al., 1985; Runyan et al., 2003), *Rattus Norvegicus* (Kwitek et al., 2004; Malfait et al., 2006). Due to its implications in *osteogenesis imperfecta* (Lee et al., 2006), *osteoporosis* (Ralston et al., 2006), *Ehlers-Danlos syndrome* (Pollitt et al, 2006), *Caffey disease* (Gensure et al., 2005), to possible implication in other diseases (e.g. *intracranial aneurysms* - Yoneyama et al., 2004), as indicator of *bone metastasis* - Fukumitsu et al., 2003) the collagen type I is studied by many researchers.

Based on information about amino acids properties and sequences on *collagen type I* chains, the aims of the study were to identify the principal property component and to analyze the similarities within and between collagens on five species.

MATERIAL AND METHOD

Amino Acids Properties

The sample of twenty essential amino acids with twenty-four associated properties has been investigated. The studied properties were: hydrophobic or hydrophilic character measured on different scales (p_1 – Black et al., 1991; p_2 – Kyte and Doolittle, 1982; p_3 - Wimley and White, 1996; p_4 – Hessa et al., 2005; p_5 – Sereda et al., 1994; p_6 - Hoop and Woods, 1981; p_7 – Cornette et al., 1987; p_8 – Eisenberg et al., 1984; p_9 – Rose et al., 1985; p_{10} – Janin, 1979; p_{11} – Engelman et al., 1986; p_{12} - Sweet and Eisenberg, 1983; p_{14} - Bull and

Breese, 1974; p_{17} – Rosema, 1988; p_{18} – Welling et al., 1985; p_{19} – Parker et al., 1986; p_{20} - Cowan and Whittaker, 1990; p_{21} - Manavalan and Ponnuswamy, 1978; p_{22} - Fauchere and Pliska, 1983; p_{23} - Rao and Argos, 1986; p_{24} – Wilson et al., 1981), partition coefficients (p_{13} - Abraham and Leo, 1987) and energies (p_{15} - Guy, 1985), inter-residue contact energies (p_{16} - Miyazawa and Jernigen, 1985). The data are presented in Table 1.

Table 1

Twenty Amino Acids: Investigated Properties

Name	L	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	p_{17}	p_{18}	p_{19}	p_{20}	p_{21}	p_{22}	p_{23}	p_{24}
Alanine	A	0.62	1.8	-0.17	0.11	47	-0.50	0.20	0.62	0.74	0.30	1.60	-0.40	0.44	0.61	0.10	5.33	0.39	1.15	2.10	0.35	12.97	0.31	1.36	-0.30
Arginine	R	0.00	-4.5	-0.81	2.58	-26	3.00	1.40	-2.53	0.64	-1.40	-12.3	-0.59	-2.42	0.69	1.91	4.18	-3.95	0.58	4.20	-1.50	11.72	-1.01	0.15	-1.10
Asparagine	N	0.24	-3.5	-0.42	2.05	-41	0.20	-0.50	-0.78	0.63	-0.50	-4.80	-0.92	-1.32	0.89	0.48	3.71	-1.91	-0.77	7.00	-0.99	11.42	-0.60	0.33	-0.20
Aspartate	D	0.03	-3.5	-1.23	3.49	-18	3.00	-3.10	-0.90	0.62	-0.60	-9.20	-1.31	-0.31	0.61	0.78	3.59	-3.81	0.65	10.0	-2.15	10.85	-0.77	0.11	-1.40
Cysteine	C	0.68	2.5	0.24	-0.13	52	-1.00	4.10	0.29	0.91	0.90	2.00	0.17	0.58	0.36	-1.42	7.93	0.25	-1.20	1.40	0.76	14.63	1.54	1.27	6.30
Glutamine	Q	0.25	-3.5	-0.58	2.36	-18	0.20	-2.80	-0.85	0.62	-0.70	-4.10	-0.91	-0.71	0.97	0.95	3.87	-1.30	-0.11	6.00	-0.93	11.76	-0.22	0.33	-0.20
Glutamate	E	0.04	-3.5	-2.02	2.68	8	3.00	-1.80	-0.74	0.62	-0.70	-8.20	-1.22	-0.34	0.51	0.83	3.65	-2.91	-0.71	7.80	-1.95	11.89	-0.64	0.25	0.00
Glycine	G	0.50	-0.4	-0.01	0.74	0	0.00	0.00	0.48	0.72	0.30	1.00	-0.67	0.00	0.81	0.33	4.48	0.00	-1.84	5.70	0.00	12.43	0.00	1.09	1.20
Histidine	H	0.16	-3.2	-0.96	2.06	-42	-0.50	0.50	-0.40	0.78	-0.10	-3.00	-0.64	-0.01	0.69	-0.50	5.10	-0.64	3.12	2.10	-0.65	12.16	0.13	0.68	-1.30
Isoleucine	I	0.94	4.5	0.31	-0.6	100	-1.80	4.80	1.38	0.88	0.70	3.10	1.25	2.46	-1.45	-1.13	8.83	1.82	-2.92	-8.00	1.83	15.67	1.80	1.44	4.30
Leucine	L	0.94	3.8	0.56	-0.55	100	-1.80	5.70	1.06	0.85	0.50	2.80	1.22	2.46	-1.65	-1.18	8.47	1.82	0.75	-9.20	1.80	14.90	1.70	1.47	6.60
Lysine	K	0.28	-3.9	-0.99	2.71	-37	3.00	-3.10	-1.50	0.52	-1.80	-8.80	-0.67	-2.45	0.46	1.40	2.95	-2.77	2.06	5.70	-1.54	11.36	-0.99	0.09	-3.60
Methionine	M	0.74	1.9	0.23	-0.1	74	-1.30	4.20	0.64	0.85	0.40	3.40	1.02	1.10	-0.66	-1.59	8.95	0.96	-3.85	-4.20	1.10	14.39	1.23	1.42	2.50
Phenylalanine	F	1.00	2.8	1.13	-0.32	92	-2.50	4.40	1.19	0.88	0.50	3.70	1.92	2.54	-1.52	-2.12	9.03	2.27	-1.41	-9.20	1.69	14.00	1.79	1.57	7.50
Proline	P	0.71	-1.6	-0.45	2.23	-46	0.00	-2.20	0.12	0.64	-0.30	-0.20	0.49	1.29	-0.17	0.73	3.87	0.99	-0.53	2.10	0.84	11.37	0.72	0.54	2.20
Serine	S	0.36	-0.8	-0.13	0.84	-7	0.30	-0.50	-0.18	0.66	-0.10	0.60	0.55	-0.84	0.42	0.52	4.09	-1.24	-0.26	6.50	-0.63	11.23	-0.04	0.97	-0.60
Threonine	T	0.45	-0.7	-0.14	0.52	13	-0.40	-1.90	-0.05	0.70	-0.20	1.20	-0.28	-0.41	0.29	0.07	4.49	-1.00	-0.45	5.20	-0.27	11.69	0.26	1.08	-2.20
Tryptophan	W	0.88	-0.9	1.85	0.3	84	-3.40	1.00	0.81	0.85	0.30	1.90	0.50	2.56	-1.20	-0.51	7.66	2.13	-1.14	-10.0	1.35	13.93	2.25	1.00	7.90
Tyrosine	Y	0.88	-1.3	0.94	0.68	49	-2.30	3.20	0.26	0.76	-0.40	-0.70	1.67	1.63	-1.43	-0.21	5.89	1.47	0.13	-1.90	0.39	13.42	0.96	0.83	7.10
Valine	V	0.83	4.2	-0.07	-0.31	79	-1.50	4.70	1.08	0.86	0.60	2.60	0.91	1.73	-0.75	-1.27	7.63	1.30	-0.13	-3.70	1.32	15.71	1.22	1.37	5.90

Collagen Type I

The alpha 1 (α1) and alpha 2 (α2) chains of collagen type I (CTI) of five species were collected from the Nat. Center for Biotechnology Information [http://www.ncbi.nlm.nih.gov/] and have been investigated: *Rattus norvegicus* (Orjel at al., 2006); *Bos taurus* (Fietzek et al., 1975; Shirai et al., 1998); *Danio rerio* (Dubois et al., 2002; Howden, 2007); *Canis lupus* (Lowe et al., 2003); and *Homo sapiens* (Simon et al., 1997; Strausberg et al., 2002).

Method of Analysis

A multivariate analysis on collagen type I amino acids properties were performed by using Principal Component Analysis (PCA) technique. A similarity analysis by using the Oliver algorithm implemented in PHP (Hypertext Preprocessor) was applied on collagen type I chains. The investigation of the minimal number of characters needed to replace, insert or delete in order to transform one amino acid chain into another has also been performed in order to analyze the similarity between and within specie.

RESULTS AND DISCUSSIONS

Principal Component on Amino Acids Properties

The following results were obtained using PCA of amino acids properties:

÷ With five exceptions the hydrophobic or hydrophilic character measured on different scales were close to each other in terms of means and standard errors. The exceptions were: p_5 (Sereda et al., 1994), p_{11} (Engelman et al., 1986), p_{19} (Parker et al., 1986), p_{21} (Manavalan and Ponnuswamy, 1978), and p_{24} (Wilson et al., 1981).

÷ The correlation matrix on investigated properties identified a very good correlation (≥ 0.95) between the following pairs of hydrophobicity: p_1 (Black et al., 1991) - p_{17} (Rosema, 1988), p_1 (Black et al., 1991) - p_{20} (Cowan and Whittaker, 1990); and p_{17} (Rosema, 1988) - p_{20} (Cowan and Whittaker, 1990).

÷ Two variables (hydrophobic or hydrophilic character scales) followed a normal distribution: p_3 (Wimley and White, 1996), and p_{18} (Welling et al., 1985).

÷ A total number of nineteen factors have been identified. The amino acids contribution on factors based on correlation matrix, expressed as extreme values (maximum and minimum) was: p_9 (Rose et al., 1985) with two maximum and one minimum, p_{13} (Abraham and Leo, 1987) with three minimum, p_{18} (Welling et al., 1985) with one maximum and three minimum.

Amino Acids on Collagen Type I

The frequency of apparition of amino acids in the alpha 1 and 2 chains according with investigated specie is presented in Table 2.

Table 2

Amino acids distribution in collagen chains

Amino acid (aa)		***Homo sapiens***		***Bos taurus***		***Canis lupus***		***Rattus norvegicus***		***Danio rerio***	
Name	Abb	HS_α1	HS_α2	BT_α1	BT_α2	CL_α1	CL_α2	TN_α1	RN_α2	DR_α1	DR_α2
Alanine	A	115	130	143	126	138	123	125	108	162	137
Cysteine	C	10	9	18	9	18	9	**0**	**0**	18	8
Aspartate	D	41	43	64	43	64	43	34	23	62	46
Glutamate	E	57	66	76	64	74	65	54	48	82	56
Phenylalanine	F	14	22	24	23	25	21	13	14	28	21
Glycine	G	**329**	**381**	**389**	**380**	**390**	**381**	**344**	**345**	**382**	**382**
Histidine	H	3	15	9	12	8	12	3	7	10	11
Isoleucine	I	10	32	25	35	26	34	8	19	36	30
Lysine	K	38	50	57	50	56	50	35	21	58	50
Leucine	L	29	61	50	60	47	59	21	34	37	57
Methionine	M	8	10	13	9	15	11	8	5	25	17
Asparagine	N	14	41	29	43	31	42	12	21	35	44
Proline	P	230	231	279	236	278	235	126	113	235	223
Glutamine	Q	30	33	51	36	50	35	25	23	40	37
Arginine	R	51	72	70	73	71	73	52	55	68	72
Serine	S	35	52	58	54	58	52	39	27	62	67
Threonine	T	23	42	44	43	47	51	16	21	53	40
Valine	V	28	55	42	50	42	50	18	39	39	35
Tryptophan	W	1	5	6	5	6	5	**0**	**0**	6	5
Tyrosine	Y	3	16	16	13	16	15	5	1	9	14
Unspecified or unknown	X	0	0	0	0	0	0	**116**	**102**	0	0
	Total	*1069*	*1366*	*1463*	*1364*	*1460*	*1366*	*1054*	*1026*	*1447*	*1352*

Looking at the data from Table 2 it can be observed that without any exceptions the glycine is the amino acid most abundant in both chains and to all investigated species. Analyzing the amino acids distribution on α1 comparing with α2 within group, statistical significant differences was been identified in two species: *Homo sapiens* (Proline, $p = 0.0019$), and *Danie rerio* (Glutamate, $p = 0.0157$). Note that the *Rattus norvegicus* specie is the one with unspecified or unknown amino acids (11% for α1 chain and almost 10% form α2 chains) and without any identified Cysteine or Tryptophan amino acids in sequences of α1 or α2 chain. The strings similarity matrix on collagen chains, expressed as percent of similarity in amino acids within and between species, was calculated (see Table 3). Following are true regarding the string similarities between species: ▪ a good similarity on α1 chain exists between *Bos taurus* - *Canis lupus*, *Homo sapiens* - *Bos Taurus*, *Bos taurus* - *Danio rerio*, and *Canis lupus*- *Danio rerio*; ▪ a good similarity on α2 chain exists between *Bos taurus* - *Canis lupus*, *Homo sapiens* - *Canis lupus*, and *Homo sapiens* - *Bos taurus*.

Table 3.

Collagen chains similarity matrix (expressed as percent)

	BT_α1	BT_α2	CL_α1	CL_α2	DR_α1	DR_α2	HS_α1	HS_α2	RN_α1
BT_α1 = *Bos taurus* CTIα1	100								
BT_α2 = *Bos taurus* CTIα2	47.19	100							
CL_α1 = *Canis lupus* CTIα1	**97.43**	47.38	100						
CL_α2 = *Canis lupus* CTIα2	48.71	**94.51**	50.18	100					
DR_α1 = *Danio rerio* CTIα1	**75.74**	49.38	**76.37**	47.00	100				
DR_α2 = *Danio rerio* CTIα2	42.77	60.16	43.46	58.13	36.73	100			
HS_α1 = *Homo sapiens* CTIα1	**82.39**	35.35	**82.56**	34.58	62.56	38.08	100		
HS_α2 = *Homo sapiens* CTIα2	53.23	**93.33**	52.09	**94.51**	49.41	66.15	40.49	100	
RN_α1 = *Rattus norvegicus* CTIα1	72.79	20.93	72.32	20.58	44.22	22.11	74.05	9.34	100
RN_α2= *Rattus norvegicus* CTIα2	18.64	70.63	14.00	66.56	25.80	40.62	15.18	66.97	38.94

Table 4.

Collagen chains Levenshtein distances within and between species

	BT_α1	BT_α2	CL_α1	CL_α2	DR_α1	DR_α2	HS_α1	HS_α2	RN_α1
BT_α1	0								
BT_α2	903	0							
CL_α1	**70**	892	0						
CL_α2	893	**98**	891	0					
DR_α1	469	902	440	896	0				
DR_α2	909	463	903	492	909	0			
HS_α1	**429**	882	448	883	731	871	0		
HS_α2	884	**115**	883	**76**	895	487	874	0	
RN_α1	**1031**	921	**1033**	921	**1016**	890	650	923	0
RN_α2	**1061**	941	**1059**	944	**1044**	923	671	940	428

The Levenshtein distances, defined as the minimal numbers of characters that must be replace, insert or delete to transform one chain into other are presented in Table 4. With seventy amino acids characters replace, insert or delete the α1 collagen type I chain of *Bos taurus* could be transformed in α1 collagen type I chain of *Canis lupus*. With ninety-eight amino acids characters replace, insert or delete the α2 collagen type I chain of *Bos taurus* could be transformed in α2 collagen type I chain of *Canis lupus*. The greater dissimilarity was identified when collagen type I chains of *Rattus norvegicus* are compared with other species (see Table 3 and 4). This could be explained by the existence of a great number of unknown or unidentified amino acids and/or by the absence of the cysteine and tryptophan (see Table 1).

CONCLUSIONS

The scale of amino acids hydrophobic or hydrophilic is still on debate even if there were proposed many methods. The scale proposed by Welling et al. seems to be more stable on principal component analysis on investigated scales.

The similarity analysis revealed that there are good similarities on Collagen type I amino acids chains between *Bos Taurus* and *Canis lupus*, as well as between *Bos Taurus* and *Homo sapiens*. Due to the lack of amino acids identification on collagen type I of *Rattus norvegicus* the greatest dissimilarities were observed when this specie was compared with the other analyzed species.

ACKNOWLEDGEMENTS

The research was partly supported by CNCSIS Romania through project AT93/2007 and UEFISCSU Romania through project ET46/2006.

BYBLIOGRAPY

1. Abraham DJ, AJ Leo, 1987, Extension of the fragment method to calculate amino acid zwitterions and side chain partition coefficients, Proteins, 2, 130-152.
2. Black SD, DR Mould, SD Black, DR Mould, 1991, Development of Hydrophobicity Parameters to Analyze Proteins Which Bear Post- or Cotranslational Modifications, Anal Biochem, 193, 72-82.
3. Bull HB, K Breese, 1974, Surface tension of amino acid solutions: A hydrophobicity scale of the amino acid residues, Arch Biochem Biophys, 161, 665-670.
4. Chu ML, W de Wet, M Bernard, F Ramirez, 1985, Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter structure, AluI repeats, and polymorphic transcripts, J Biol Chem, 260(4), 2315-2320.
5. Cornette JL, KB Cease, H Margalit, JL Spouge, JA Berzofsky, C DeLisi, 1987, Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins, J Mol Biol, 195, 659-685.
6. Cowan R, RG Whittaker, 1990, Hydrophobicity indices for amino acid residues as determined by HPLC, Pept Res, 3, 75-80.
7. Dubois GM, Z Haftek, C Crozet, R Garrone, D Le Guellec, 2002, Structure and spatio temporal expression of the full length DNA complementary to RNA coding for alpha2 type I collagen of zebrafish, Gene, 294, 55-65.
8. Eisenberg D, E Schwarz, M Komaromy, R Wall, 1984, Analysis of membrane and surface protein sequences with the hydrophobic moment plot, J Mol Biol, 179, 125-142.
9. Engelman DM, TA Steitz, A Goldman, 1986, Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins, Annu Rev Biophys Biophys Chem, 15, 321-353.
10. Fauchere J-L, VE Pliska, 1983, Hydrophobic parameters of amino-acid side-chains from the partitioning of N-acetyl-amino-acid amide, Eur J Med Chem, 18, 369-375.
11. Fietzek PP, K Kuhn, 1975, The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide peptides alpha1-CB2, alpha1-CB4 and alpha1-CB5 from calf-skin collagen, Eur J Biochem, 52 (1), 77-82.
12. Fukumitsu N, M Uchiyama, Y Mori, K Kishimoto, J Nakada, 2003, A comparative study of prostate specific antigen (PSA), C-terminal propeptide of blood type I procollagen (PICP) and urine type I collagen-crosslinked N telopeptide (NTx) levels using bone scintigraphy in prostate cancer patients, Ann Nucl Med, 17(4), 297-303.
13. RC Gensure, O Mäkitie, Catherine Barclay, Catherine Chan, SR DePalma, M Bastepe, H Abuzahra, R Couper, S Mundlos, D Sillence, Leena Ala Kokko, JG Seidman, WG Cole, H Jüppner, 2005, A novel Col1A1 mutation in infantile cortical hyperostosis (Caffey disease) expands the spectrum of collagen-related disorders, J Clin Invest, 115(5), 1250-1257.
14. Guy HR, 1985, Amino acid side-chain partition energies and distribution of residues in soluble proteins, Biophys J, 47, 61-70.
15. Hessa Tara, H Kim, K Bihlmaier, Carolina Lundin, Jorrit Boekel, Helena Andersson, IngMarie Nilsson, SH White and G von Heijne, 2005, Recognition of transmembrane helices by the endoplasmic reticulum translocon, Nature, 433, 377-381.
16. Hoop TP, KR Woods, 1981, Prediction of protein antigenic determinants from amino acid sequences, Proc Natl Acad Sci USA, 78, 3824-3828.
17. Howden P, Direct Submission. Submitted (05-JUN-2007) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/Projects/D_rerio/faqs.shtml#dataeight
18. Janin J, 1979, Surface and inside volumes in globular proteins, Nature, 277, 491-492.
19. Kadler KE, DF Holmes, JA Trotter, JA Chapman, 1996, Collagen fibril formation, Biochem J, 316, 1-11.
20. Kwitek Anne E., Jo Gullings-Handley, J Yu, DC Carlos, Kimberly Orlebeke, J Nie,1 J Eckert, Angela Lemke, Jaime Wendt Andrae, Susan Bromberg, D Pasko, D Chen, TE Scheetz, TL Casavant, MB Soares, VC Sheffield, PJ Tonellato, HJ Jacob, 2004, High-density rat radiation hybrid maps containing over 24,000 SSLPs, genes, and ESTs provide a direct link to the rat genome sequence, Genome Res, 14(4), 750-757.
21. Kyte J, RF Doolittle, 1982, A Simple Method for Displaying the Hydropathic Character of a Protein, J Mol Biol, 157, 105-132.
22. Lee K-S, H-R Song, T-J Cho, HJ Kim, T-M Lee, H-S Jin, H-Y Park, S Kang, S-C Jung, SK Koo, 2006, Mutational spectrum of type I collagen genes in Korean patients with osteogenesis imperfecta, Hum Mutat, 27(6):599.
23. Lowe JK, R Guyon, ML Cox, DC Mitchell, AL Lonkar, F Lingaas, C André, F Galibert, EA Ostrander, KE Murphy, 2003, Radiation hybrid mapping of the canine type I and type IV collagen gene subfamilies, Funct Integr Genomics, 3, 112-116.
24. Malfait F, S Symoens, P Coucke, L Nunes, S De Almeida, A De Paepe, 2006, Total absence of the alpha2(I) chain of collagen type I causes a rare form of Ehlers-Danlos syndrome with hypermobility and propensity to cardiac valvular problems, J Med Genet, 43(7), E36.

25. Manavalan P, PK Ponnuswamy, 1978, Hydrophobic character of amino acid residues in globular proteins, Nature, 275, 673-674.
26. Miyazawa S, RL Jernigen, 1985, Estimation of Effective Interresidue Contact Energies from Protein Crystal Structures: Quasi-Chemical Approximation, Macromolecules, 18, 534-552.
27. Orgel JP, TC Irving, A Miller, TJ Wess, 2006, Microfibrillar structure of type I collagen in situ, Proc Natl Acad Sci USA, 103(24), 9001-9005.
28. Parker JMR, D Guo, RS Hodges, 1986, New Hydrophilicity Scale Derived from High-Performance Liquid Chromatography Peptide Retention Data: Correlation of Predicted Surface Residues with Antigenicity and X-ray-Derived Accessible Sites?, Biochemistry, 25, 5425-5431.
29. Pollitt R, R McMahon, J Nunn, R Bamford, A Afifi, N Bishop, A Dalton, 2006, Mutation analysis of Col1A1 and Col1A2 in patients diagnosed with osteogenesis imperfecta type I-IV, Hum Mutat, 27(7), 716.
30. Ralston SH, AG Uitterlinden, Maria L Brandi, Susana Balcells, BL Langdah, P Lips, R Lorenc, Barbara Obermayer-Pietsch, Serena Scollen, Mariona Bustamante, Lise Bjerre Husted, AH Carey, A Diez-Perez, Alison M Dunning, A Falchetti, Elzbieta Karczmarewicz, M Kruk, JPTM van Leeuwen, JBJ van Meurs, J Mangion, Fiona EA McGuigan, L Mellibovsky, Francesca del Monte, HAP Pols, J Reeve, DM Reid, W Renner, F Rivadeneira, Natasja M. van Schoor, RE Sherlock, JPA Ioannidis, 2006, Large-scale evidence for the effect of the ColIA1 Sp1 polymorphism on osteoporosis outcomes: the GENOMOS study, PLoS Med, 3(4):e90.
31. Rao MJK, P Argos, 1986, A conformational preference parameter to predict helices in integral membrane proteins, Biochim Biophys Acta, 869, 197-214.
32. Rose GD, AR Geselowitz, GJ Lesser, RH Lee, MH Zehfus, 1985, Hydrophobicity of amino acid residues in globular proteins, Science, 229, 834-838.
33. Roseman MA, 1988, Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds, J Mol Biol, 200, 513-522.
34. Runyan CE, HW Schnaper, AC Poncelet, 2003, Smad3 and PKCdelta mediate TGF-beta1-induced collagen I expression in human mesangial cells, Am J Physiol Renal Physiol, 285(3), F413-F422.
35. Sereda TJ, CT Mant, FD Sonnichsen, RS Hodges, 1994, Reversed-phase chromatography of synthetic amphipathic a-helical peptides as a model for ligand/receptor interactions effect of changing hydrophobic environment on the relative hydrophilicity/hydrophobicity of amino acid side- chains, J Chromatogr A, 676(1), 139-153.
36. Shirai T, S Hattori, M Sakaguchi, S Inouye, A Kimura, T Ebihara, S Irie, Y Nagai, H Hori, 1998, The complete cDNA coding sequence for the bovine proalpha2(I) chain of type I procollagen, Matrix Biol, 17 (1), 85-88.
37. Simon MP, F Pedeutour, N Sirvent, J Grosgeorge, F Minoletti, JM Coindre, MJ Terrier-Lacombe, N Mandahl, RD Craver, N Blin, G Sozzi, C Turc-Carel, KP O'Brien, D Kedra, I Fransson, C Guilbaud, JP Dumanski, 1997, Deregulation of the platelet-derived growth factor B-chain gene via fusion with collagen gene Col1A1 in dermatofibrosarcoma protuberans and giant-cell fibroblastoma, Nat Genet, 15, 95-98.
38. Strausberg RL, Elise A Feingold, Lynette H Grouse, JG Derge, RD Klausner, FS Collins, L Wagner, Carolyn M Shenmen, GD Schuler, SF Altschul, B Zeeberg, Kenneth H Buetow, CF Schaefer, Narayan K Bhat, RF Hopkins, Heather Jordan, Troy Moore, SI Max, J Wang, Florence Hsieh, Luda Diatchenko, Kate Marusina, AA Farmer, GM Rubin, L Hong, M Stapleton, MB Soares, Maria F Bonaldo, TL Casavant, TE Scheetz, MJ Brownstein, TB Usdin, S Toshiyuki, P Carninci, Christa Prange, SS Raha, Naomi A Loquellano, GJ Peters, RD Abramson, Sara J Mullahy, Stephanie A Bosak, PJ McEwan, KJ McKernan, Joel A Malek, Preethi H Gunaratne, S Richards, Kim C Worley, Sarah Hale, Angela M Garcia, Laura J Gay, SW Hulyk, Debbie K Villalon, Donna M Muzny, Erica J Sodergren, Xiuhua Lu, RA Gibbs, Jessica Fahey, E Helton, M Ketteman, M Madan, Stephanie Rodrigues, Amy Sanchez, Michelle Whiting, A Madan, Alice C Young, Y Shevchenko, GG Bouffard, RW Blakesley, JW Touchman, ED Green, MC Dickson, AC Rodriguez, Jane Grimwood, J Schmutz, RM Myers, YSN Butterfield, MI Krzywinski, Ursula Skalska, Duane E Smailus, Angelique Schnerch, Jacqueline E Schein, SJM Jones, MA Marra and Mammalian Gene Collection (MGC) Program Team, 2002, Generation and initial analysis of more than 15,000 full-lengthhuman and mouse cDNA sequences, Proc Natl Acad Sci USA, 99(26), 16899-16903.
39. Sweet RM, D Eisenberg, 1983, Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure, J Mol Biol, 171, 479-488.
40. Veit G, B Kobbe, DR Keene, M Paulsson, M Koch, R Wagener, 2006, Collagen XXVIII, a novel von Willebrand factor A domain-containing protein with many imperfections in the collagenous domain, J Biol Chem, 281(6), 3494-504.
41. Welling GW, WJ Weijer, R Van der Zee, S Welling-Wester, 1985, Prediction of sequential antigenic regions in proteins, FEBS Lett, 188, 215-218.
42. Wilson KJ, A Honegger, RP Stotzel, GJ Hughes, 1981, The behaviour of peptides on reverse-phase supports during high-pressure liquid chromatography, Biochem J, 199, 31-41.
43. Wimley WC, SH White, 1996, Experimentally determined hydrophobicity scale for proteins at membrane interfaces, Nature Struct Biol, 3, 842-848.

Bulletin of the University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca
3 Manastur Street
Cluj-Napoca 400372 România

Edited by AcademicPres®

EAP EDITURA ACADEMICPRES