Crystal Structure and Molecular Properties of Lycopene Similarities for Ionizing Radiation Protection

Jeff Cromwell, PhD

The Mathematical Learning Space Research Portfolio

1 Abstract

Lycopene is a linear, unsaturated hydrocarbon carotenoid used to reduce the frequency or rate of spontaneous or induced tumors independently of the mechanism involved and it can be a Radiation-Protective Agents Drugs used to protect against ionizing radiation in radiation therapy. In addition, it can be used as Anti-Inflammatory Agents to inhibit inflammation with reduction and as antioxidants that inhibit oxidation reactions.A collection of 50 proteins similar to the Lycopene 1VHQ crystal were examined with standard molecular properties and compared and contrasted. The modes and fluctuations of the lycopene crystal was examined and a cluster analysis performed of their descriptive statistics with an emphasis on the third moment. In addition, comparisions were made with the overall structural change measured by the radius of gyration for all the crystals.

2 Introduction

Lycopene is a linear, unsaturated hydrocarbon carotenoid with a reduction in the frequency or rate of spontaneous or induced tumors independently of the mechanism involved or radiation-protective agents drugs used to protect against ionizing radiation and used in radiation therapy. In addition, it is an anti-inflammatory agent that reduces and suppresses inflammation along with antioxidants that inhibit oxidation reactions. [800]

Based on the collection of PMIDs for CID 446925 Table 1 has the PMIDs and Titles from an abstract search with the keyword RADIATION. [601]

*

Email address: `http://mathlearningspace.weebly.com/ @` (Jeff Cromwell, PhD)

ID	PMID	Title
21	22549412	Alleviation of chilling injury in postharvest tomato fruit by preconditioning with ultraviolet irradiation
40	22692080	Decylglucoside-based microemulsions for cutaneous localization of lycopene and ascorbic acid
55	22743816	Nutrition support and therapy in patients with head and neck squamous cell carcinomas
275	21294875	The use of some nanoemulsions based on aqueous propolis and lycopene extract in the skin's protective mechanisms against UVA radiation
285	20853276	The effect of environmental conditions on nutritional quality of cherry tomato fruits: evaluation of two experimental Mediterranean greenhouses
286	20151398	Determination of the influence of IR radiation on the antioxidative network of the human skin
315	20637178	Significant correlations of dermal total carotenoids and dermal lycopene with their respective plasma levels in healthy adults
364	22254062	The role of phytonutrients in skin health
374	20041432	Lycopene protects the structure of the small intestine against gamma-radiation-induced oxidative stress
419	19591120	A simple and rapid method to assess lycopene in multiple layers of skin samples
427	19681037	Resonance Raman spectroscopy as an effective tool for the determination of antioxidative stability of cosmetic formulations
432	20090407	Radical production by infrared A irradiation in human tissue
514	19450652	Lycopene: An antioxidant and radioprotector against gamma-radiation-induced cellular damages in cultured human lymphocytes
610	20155616	Lycopene in the prevention of gastrointestinal toxicity of radiotherapy
627	18959418	Influence of electron-beam irradiation on bioactive compounds in grapefruits (Citrus paradisi Macf.)
645	18696411	Phytochemicals as protectors against ultraviolet radiation: versatility of effects and mechanisms
819	17368731	Evaluation of the antioxidant effects of carotenoids from Deinococcus radiodurans through targeted mutagenesis, chemiluminescence, and DNA damage analyses
838	17189673	Lycopene as a natural protector against gamma-radiation induced DNA damage, lipid peroxidation and antioxidant status in primary culture of isolated rat hepatocytes in vitro
888	17992006	Culture characteristics of carotenoid-producing filamentous fungus T-1, and carotenoid production
986	16736578	[Construction and functional analysis of the crtl gene disruptant in Deinococcus radiodurans]
1017	16689237	[Combined effects of enhanced UV-B radiation and doubled CO2 on tomato growth and its fruit quality]
1080	16018824	Prostate cancer: chemoprevention update 2005
1086	15884827	Bioactive compounds of grapefruit (Citrus paradisi Cv. Rio Red) respond differently to postharvest irradiation, storage, and freeze drying
1255	14678532	Antioxidant activity of topically applied lycopene
1570	11265592	Role of lycopene in recovery of radiation induced injury to mammalian cellular organelles

Figure 1 has the crystal structure of 1VHQ enhancing lycopene biosynthesis protein 2 from one of a collection of N=50 proteins from the structural analysis of a set of proteins resulting from a bacterial genomics project [1]

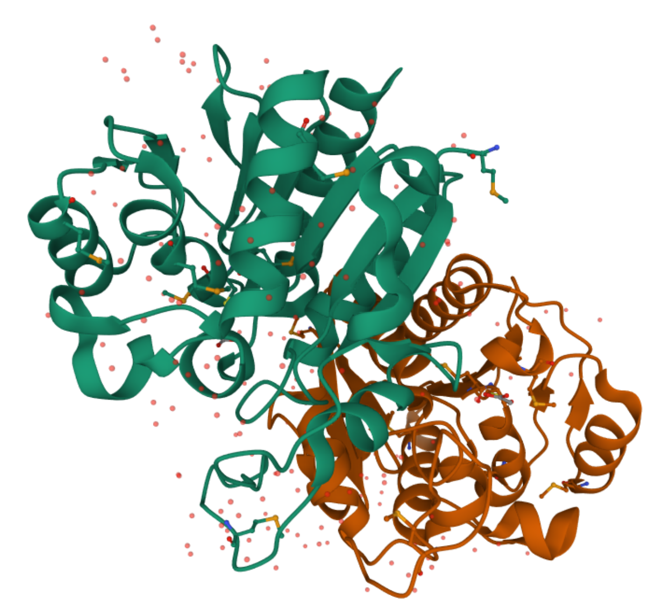

Figure 1: Crystal structure of enhancing lycopene biosynthesis protein 2 1VHQ [1] [2] [1002]

Based on the collection of PMIDs for CID 446925 Table 2 has the PMIDs and Titles from an abstract search with the keyword BACTERIA. [601]

ID	PMID	Title
38	23112480	Counteraction of reactive oxygen species and determination of antibacterial efficacy of proanthocyanidin and lycopene when mixed with calcium hydroxide and chlorhexidine mixture: An in vitro comparative study
77	21705442	Phaeospirillum tilakii sp. nov., a phototrophic alphaproteobacterium isolated from aquatic sediments
81	22070388	Biosynthesis of the biomarker okenone: Ï‡-ring formation
155	22745782	On the structure and function of the phytoene desaturase CRTI from Pantoea ananatis, a membrane-peripheral and FAD-dependent oxidase/isomerase
171	21921032	Elucidation of the biosynthetic pathway for Okenone in Thiodictyon sp. CAD16 leads to the discovery of two novel carotene ketolases
175	21663487	Phenolic composition and antimicrobial and antioxidant activities of Leucoagaricus leucothites (Vittad.) Wasser
212	21590288	Biochemical characterization of the carotenoid 1,2-hydratases (CrtC) from Rubrivivax gelatinosus and Thiocapsa roseopersicina
214	20709914	Phaeospirillum oryzae sp. nov., a spheroplast-forming, phototrophic alphaproteobacterium from a paddy soil
218	21547544	Genus specific unusual carotenoids in purple bacteria, Phaeospirillum and Roseospira: structures and biosyntheses
263	21209101	Plant carotene cis-trans isomerase CRTISO: a new member of the FAD(RED)-dependent flavoproteins catalyzing non-redox reactions
378	20629863	Chemical characterization of tomato juice fermented with bifidobacteria
505	19481130	Antimicrobial and antioxidant activities of Russula delica Fr
515	19514731	Tomatine-containing green tomato extracts inhibit growth of human breast, colon, liver, and stomach cancer cells
539	19304845	The biosynthetic pathway for myxol-2' fucoside (myxoxanthophyll) in the cyanobacterium Synechococcus sp. strain PCC 7002
550	19329583	Ectothiorhodospira variabilis sp. nov., an alkaliphilic and halophilic purple sulfur bacterium from soda lakes
554	18948076	Evolution of carotene desaturation: the complication of a simple pathway
571	18930627	Modulation of experimental osteoporosis in rats by the antioxidant beverage effective microorganism-X (EM-X)
644	18676669	Isorenieratene biosynthesis in green sulfur bacteria requires the cooperative actions of two carotenoid cyclases
690	18417238	Carotenoid accumulation in bacteria with enhanced supply of isoprenoid precursors by upregulation of exogenous or endogenous pathways
719	17920247	Synergistic effect between lycopene and ciprofloxacin on a chronic bacterial prostatitis rat model
820	17434985	Maize Y9 encodes a product essential for 15-cis-zeta-carotene isomerization
827	17428435	Kinetic variations determine the product pattern of phytoene desaturase from Rubrivivax gelatinosus
855	17896581	The effect of daily consumption of probiotic and conventional yoghurt on oxidant and anti-oxidant parameters in plasma of young healthy women
934	16967101	Carotenoid biosynthesis in cyanobacteria: structural and evolutionary scenarios based on comparative genomics
953	16625353	A carotenoid synthesis gene cluster from Algoriphagus sp. KK10202C with a novel fusion-type lycopene beta-cyclase gene
986	16736578	[Construction and functional analysis of the crtI gene disruptant in Deinococcus radiodurans]
1031	16235086	Carotenoid biosynthesis in Gloeobacter violaceus PCC4721 involves a single crtI-type phytoene desaturase instead of typical cyanobacterial enzymes
1079	15933032	Novel carotenoid oxidase involved in biosynthesis of 4,4'-diapolycopene dialdehyde
1092	15821145	Why is golden rice golden (yellow) instead of red?
1096	16119846	[Isoprenoid pigments in representatives of the family Microbacteriaceae]
1101	15811329	The cyanobacterium Gloeobacter violaceus PCC 7421 uses bacterial-type phytoene desaturase in carotenoid biosynthesis
1123	15716108	Carotenoid biosynthetic pathway: molecular phylogenies and evolutionary behavior of crt genes in eubacteria
1230	14734565	Two distinct crt gene clusters for two different functional classes of carotenoid in Bradyrhizobium
1300	14642353	Coordinate expression of multiple bacterial carotenoid genes in canola leading to altered carotenoid production
1328	12763034	Fusion-type lycopene beta-cyclase from a thermoacidophilic archaeon Sulfolobus solfataricus
1330	12712234	A novel type of lycopene epsilon-cyclase in the marine cyanobacterium Prochlorococcus marinus MED4
1360	12664151	Surrogate biochemistry: use of Escherichia coli to identify plant cDNAs that impact metabolic engineering of carotenoid accumulation
1410	12358437	Tomato glycoalkaloids: role in the plant and in the diet
1472	11884677	Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis
1502	11526111	Alteration of product specificity of Rhodobacter sphaeroides phytoene desaturase by directed evolution
1524	11724407	Gene sll0033 from Synechocystis 6803 encodes a carotene isomerase involved in the biosynthesis of all-E lycopene
1572	11357508	Dihydroxylycopene diglucoside diesters: a novel class of carotenoids from the phototrophic purple sulfur bacteria Halorhodospira abdelmalekii and Halorhodospira halochloris
1587	11387982	Biosynthesis of beta-carotene (provitamin A) in rice endosperm achieved by genetic engineering

Molecular properties are abundant in dimensional reduction for collections of sequences. Examples such as stability, binding potential aliphatic and hydrophobicity along with the charge at different pH is a few of available molecular properties to examine based on the gene ontology ids in Table 2. The net charge of a protein sequence based on the Henderson-Hasselbalch equation based on pH 5, 7 and 9. The aliphatic index is the relative volume occupied by aliphatic side chains (Alanine, Valine, Isoleucine, and Leucine) and is a positive factor for the increase of thermostability of globular proteins. The potential protein interaction index proposed by Boman (2003) based in the amino acid sequence of a protein and provides an overall estimate of the potential of a peptide to bind to membranes or other proteins as receptors. A protein have high binding potential if the index value is higher than 2.48. This index predicts the stability of a protein based on its amino acid composition, a protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable. Hydrophobicity is an important stabilization force in protein folding; this force changes depending on the solvent in which the protein is found.

Table 3 provides values based on each one of these properties.

	Name	Stability Index	Binding Potential	ALiphatic	f.1	CpH5	CpH7	CpH9
1	1O60	31.5	1.100	95.3	0.004833	7.5482	-16.690	-38.650
2	1O61	40.0	1.452	94.6	-0.268868	19.5472	0.498	-17.259
3	1O62	40.0	1.452	94.6	-0.268868	19.5472	0.498	-17.259
4	1O63	30.3	1.447	105.3	-0.059057	12.5520	0.706	-8.832
5	1O64	30.3	1.447	105.3	-0.059057	12.5520	0.706	-8.832
6	1O65	41.4	1.740	80.5	-0.404755	3.3080	-18.055	-29.893
7	1O66	35.1	0.576	89.8	0.273963	18.9543	-17.152	-50.102
8	1O67	41.9	1.724	80.9	-0.382288	-1.7731	-22.296	-33.818
9	1O68	35.7	0.586	89.0	0.262490	24.1166	-13.671	-47.290
10	1O69	40.0	1.452	94.6	-0.268868	19.5472	0.498	-17.259
11	1O6B	44.3	1.704	89.6	-0.173620	2.0256	-2.378	-6.554
12	1O6C	37.2	1.668	89.3	-0.330124	6.5574	-23.448	-33.038
13	1O6D	39.3	1.598	102.0	-0.087755	9.2268	3.169	0.172
14	1VGT	32.4	1.388	102.3	-0.046226	13.0164	-6.420	-18.367
15	1VGU	35.6	1.505	101.2	-0.070244	13.0467	-5.690	-18.042
16	1VGV	43.1	1.521	99.0	-0.149429	36.1370	-19.864	-48.321
17	1VGW	36.3	1.295	110.6	0.010055	23.5477	-8.487	-25.568
18	1VGX	32.1	1.638	86.8	-0.351667	1.6807	-12.749	-20.700
19	1VGY	34.3	1.561	85.6	-0.271200	8.7114	-17.080	-34.800
20	1VGZ	37.3	1.332	110.8	-0.004028	7.6531	-3.177	-9.343
21	1VH0	42.4	1.637	92.3	-0.392151	50.5318	29.947	22.086
22	1VH1	28.2	1.575	97.3	-0.082497	-6.9959	-38.377	-53.113
23	1VH2	31.7	1.724	86.6	-0.374342	0.9712	-6.385	-10.708
24	1VH3	29.3	1.415	105.3	-0.081583	4.2967	-15.441	-24.332
25	1VH4	40.6	1.773	92.4	-0.342476	39.0838	-3.079	-20.641
26	1VH5	41.4	1.499	97.5	-0.080435	11.1484	0.168	-9.102
27	1VH6	43.6	1.830	90.7	-0.338095	-4.1788	-7.064	-10.130
28	1VH7	23.7	1.364	103.7	0.030800	-0.3740	-7.059	-10.026
29	1VH8	37.6	1.267	104.0	0.013472	22.8254	-14.639	-34.289
30	1VH9	30.6	1.549	86.2	-0.132609	10.4250	-2.938	-14.484
31	1VHA	37.6	1.272	103.9	0.009190	23.8254	-13.639	-33.304
32	1VHC	37.0	0.672	119.9	0.305106	0.7576	-16.833	-30.275
33	1VHD	38.5	1.437	88.0	-0.246260	10.2308	-10.695	-20.514
34	1VHE	20.6	1.314	92.2	-0.153425	6.2313	-7.131	-12.346
35	1VHF	44.7	1.642	93.7	-0.409901	-0.0185	-2.798	-4.997
36	1VHG	37.4	1.772	93.2	-0.265946	-8.1403	-17.009	-19.843
37	1VHJ	25.5	1.200	87.5	0.021439	1.2968	-34.052	-68.268
38	1VHK	44.2	1.626	90.4	-0.234701	5.7104	-15.708	-37.952
39	1VHL	31.3	1.472	108.4	-0.081970	7.7993	-13.161	-19.811
40	1VHM	36.3	1.279	102.7	0.093730	-8.1606	-15.212	-22.042
41	1VHO	29.8	1.215	103.1	0.006583	2.1660	-7.397	-13.040
42	1VHQ	35.3	0.825	108.2	0.275982	-11.6837	-23.353	-33.867
43	1VHS	39.0	1.545	85.9	-0.243034	0.5504	-10.124	-15.464
44	1VHT	31.3	1.476	108.3	-0.084628	6.9540	-14.919	-21.808
45	1VHU	23.1	0.982	94.0	0.015104	1.8083	-3.373	-7.856
46	1VHV	30.0	1.408	102.5	-0.029026	11.0648	-9.812	-20.491
47	1VHW	26.0	1.232	86.6	0.000564	-0.4796	-36.050	-70.268
48	1VHX	30.8	1.218	104.5	-0.042446	0.4870	-2.996	-4.651
49	1VHY	44.5	1.484	104.1	-0.189331	13.8141	-2.970	-16.821
50	1VHZ	38.5	1.774	93.2	-0.268464	-8.1403	-17.009	-19.843

Table 3a has the AA properties with the molecular percentages from the molecules in Table 2. The Pearson Type is 1 with a=0.848497, b= 0.9896716, location=5.755 and scale=54.84263 for the Acidic Molecular Percentage given by the Beta distribution probability density function with parameters a, b, scale=s and location=λ is given by

$$f(x) = \frac{\Gamma(a+b)}{(\Gamma(a)\Gamma(b))}((x-\lambda)/s)^{(a-1)}(1-((x-\lambda)/s))^{(b-1)} \qquad (1)$$

for $a > 0$, $b > 0$, $s <> 0$, $0 < (x-\lambda)/s < 1$.

	ID	Tiny	Small	Aliphatic	Aromatic	Nonpolar	Polar	Charged	Basic	Acidic
1	1O60	26.39	44.78	48.28	24.35	31.58	14.00	9.04	11.69	60.42
2	1O61	52.05	27.76	30.95	12.57	9.87	15.60	55.34	28.21	39.58
3	1O62	32.71	13.50	11.28	11.78	55.26	30.89	44.66	52.66	28.65
4	1O63	8.55	14.26	51.99	27.33	44.74	49.29	27.40	35.74	14.06
5	1O64	57.25	37.10	48.01	51.33	30.92	37.46	13.97	10.66	14.58
6	1O65	42.75	55.66	30.12	31.67	15.13	9.20	13.43	56.74	27.24
7	1O66	24.16	37.92	14.86	10.00	15.79	59.15	19.80	43.26	47.12
8	1O67	11.90	8.22	15.27	55.67	22.16	40.85	32.67	23.82	35.19
9	1O68	12.27	63.22	20.41	44.33	49.93	27.71	29.70	10.03	10.34
10	1O69	25.47	36.78	35.37	30.00	37.27	14.57	14.85	13.79	54.27
11	1O6B	43.26	21.32	30.61	14.67	9.78	13.14	51.48	25.70	45.73
12	1O6C	31.40	10.98	12.24	15.33	55.97	34.78	48.52	51.41	31.01
13	1O6D	11.99	10.33	53.74	32.00	44.03	54.35	33.66	34.17	15.90
14	1VGT	52.43	26.96	46.26	53.33	25.90	29.71	15.84	9.40	15.11
15	1VGU	47.57	46.87	36.05	30.13	12.66	9.42	17.82	57.37	31.10
16	1VGV	28.30	28.53	20.41	9.33	13.24	55.07	22.16	42.63	54.80
17	1VGW	14.82	13.64	15.65	53.60	28.76	44.93	47.57	27.59	32.51
18	1VGX	13.48	55.64	31.60	46.40	48.30	24.64	30.81	13.48	9.73
19	1VGY	25.47	44.36	54.48	26.13	32.77	13.77	9.73	14.11	59.59
20	1VGZ	43.26	27.27	38.68	13.07	11.65	10.87	54.05	32.79	40.41
21	1VH0	31.40	12.85	9.67	13.07	50.24	30.85	45.95	54.97	25.74
22	1VH1	11.99	14.42	58.02	31.04	49.76	49.23	28.65	40.19	12.27
23	1VH2	52.43	36.88	41.98	52.61	25.73	37.42	12.43	5.54	13.47
24	1VH3	47.57	55.16	26.18	41.23	14.81	9.19	16.22	61.43	29.50
25	1VH4	28.30	37.43	14.15	5.69	10.92	59.08	30.84	38.57	48.56
26	1VH5	14.82	8.75	12.03	56.63	31.88	40.92	54.77	23.33	34.53
27	1VH6	13.48	63.20	31.71	43.37	48.55	27.79	32.91	9.70	5.75
28	1VH7	23.08	36.80	54.63	24.88	31.88	14.66	9.83	13.63	55.40
29	1VH8	45.16	21.36	38.29	12.80	8.70	13.13	60.26	29.41	44.60
30	1VH9	31.76	11.19	9.27	12.09	55.07	27.81	39.74	46.13	25.54
31	1VHA	8.93	10.16	57.32	21.72	44.93	51.38	25.86	30.34	12.23
32	1VHC	54.59	25.47	42.68	36.77	27.54	41.08	12.39	13.62	13.31
33	1VHD	45.41	43.26	27.07	29.57	15.22	6.13	13.46	56.66	27.41
34	1VHE	30.02	31.40	14.63	9.03	12.32	62.29	28.61	43.34	42.68
35	1VHF	15.88	11.99	12.44	51.18	23.81	37.71	44.90	27.86	33.89
36	1VHG	14.14	52.43	25.25	48.82	43.81	20.74	30.83	13.31	7.53
37	1VHJ	23.08	47.57	49.03	29.14	31.91	9.90	9.50	14.55	53.98
38	1VHK	45.16	28.30	33.11	16.67	8.57	10.84	52.99	27.77	46.02
39	1VHL	31.76	14.82	10.21	12.47	52.86	27.70	47.01	52.56	27.20
40	1VHM	8.93	13.48	54.80	25.93	47.14	48.75	26.96	40.99	14.44
41	1VHO	54.59	26.99	45.20	51.55	24.29	29.36	13.13	8.10	12.76
42	1VHQ	45.41	49.08	26.46	36.50	10.48	10.25	13.83	55.21	22.37
43	1VHS	30.02	30.06	13.90	9.82	13.81	53.74	27.71	44.79	47.44
44	1VHT	15.88	11.04	12.56	59.02	28.80	46.26	52.59	23.97	30.73
45	1VHU	14.14	52.15	31.66	40.98	49.60	29.92	40.90	12.07	9.70
46	1VHV	26.69	47.85	53.18	27.75	35.60	14.96	8.01	11.90	53.91
47	1VHW	46.17	25.77	41.16	12.81	8.80	14.96	55.26	30.21	46.09
48	1VHX	28.37	12.88	5.66	14.94	55.60	30.68	44.74	48.44	28.57
49	1VHY	13.65	12.88	57.11	26.97	44.40	51.78	23.71	35.94	12.40
50	1VHZ	55.22	24.62	42.89	50.66	29.60	32.60	12.02	10.42	16.17

Table 1: Tiny=(A + C + G + S + T), Small= (A + B + C + D + G + N + P + S + T + V), Aliphatic =(A + I + L + V), Aromatic= (F + H + W + Y), Non-polar=(A + C + F + G + I + L + M + P + V + W + Y), Polar=(D + E + H + K + N + Q + R + S + T + Z), Charged=(B + D + E + H + K + R + Z), Basic=(H + K + R) and Acidic=(B + D + E + Z)

A cluster analysis was performed based on the Amino Acid Molecular Percentages from Table 1A and the cluster dendrogram presented in Figure 1A.

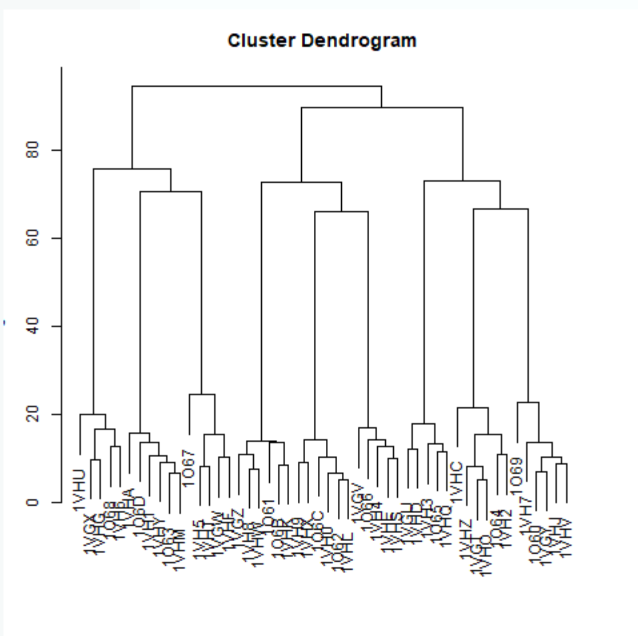

Figure 2: Cluster Analysis of Amino Acid Molecular Percentages [1001]

Figure 2 has the modes for 1VHQ with the maximum likelihood estimates of the Pearson distribution identification of type 5 with shape 1.79 location 0.04 and scale 0.19.

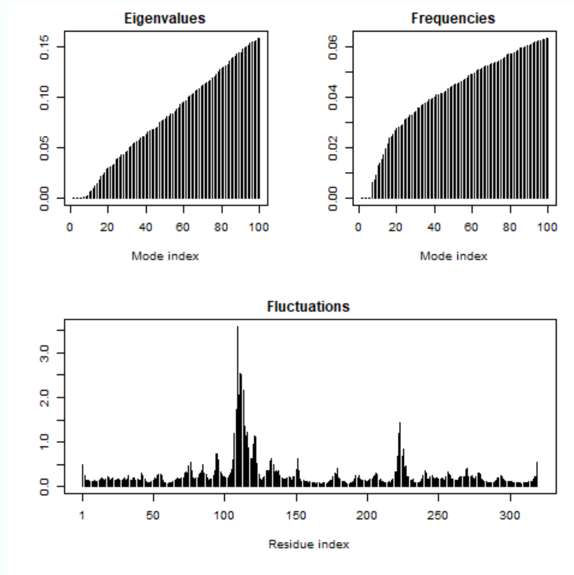

Figure 3: Mode and Fluctuations for Summary for Crystal structure of enhancing lycopene biosynthesis protein 2 1VHQ [1] [2][1002]

Based on the indicies between 100 and 150, the change in amplitude is the highest with another change in the interval 200 and 250. Table 3 has the summary statistic for all three intervals of (a) 1:319, (b) 100:150 and (c) 200:250 for 1VHQ. Since there are two 1VHQ in Table 4, a comparison was done with fluctuations with a Pearson's product-moment correlation correlation test for the first 300 modes with t = 0.79943, df = 298, p-value = 0.4247. The alternative hypothesis: true correlation is not equal to 0 with 95 percent confidence interval: -0.06733344 0.15866977 and the sample estimate correlation of 0.04626013.

	1:319	100:150	200:250
vars	1.00	1.00	1.00
n	319.00	51.00	51.00
mean	0.28	0.69	0.27
sd	0.37	0.75	0.26
median	0.18	0.35	0.17
trimmed	0.20	0.53	0.20
mad	0.09	0.25	0.09
min	0.07	0.14	0.09
max	3.59	3.59	1.42
range	3.52	3.45	1.33
skew	4.88	1.95	2.73
kurtosis	29.69	3.56	7.67
se	0.02	0.10	0.04

Figure 3 has Cluster Analysis and Dendrogram of the fluctuations of the 50 Lycopene Crystal Similarities based on the summary statistics from Table 2.

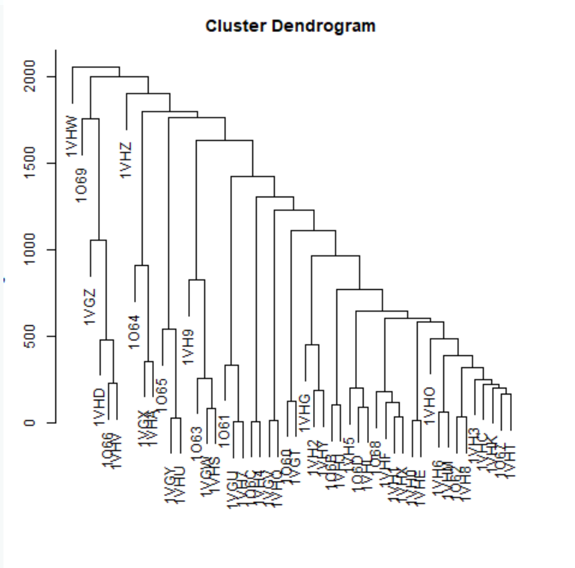

Figure 4: Cluster Analysis and Dendrogram of the Summary Statistics for the Fluctuations from 50 Lycopene Similarities [1]

Figure 4 has the empirical distribution for the skewness of the fluctuations for the 50 Lycopene Crystal Similarities.

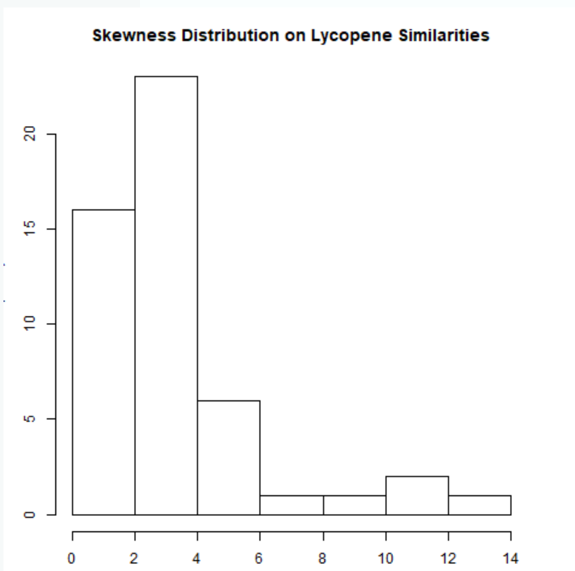

Figure 5: Distribution of Skewness of the Fluctuations for 50 Lycopene Crystal Similarities [2][1002]

Figure 5 has the empirical distribution of Radius of gyration for the overall structural change of macromolecules for the 50 Lycopene Crystal Similarities.

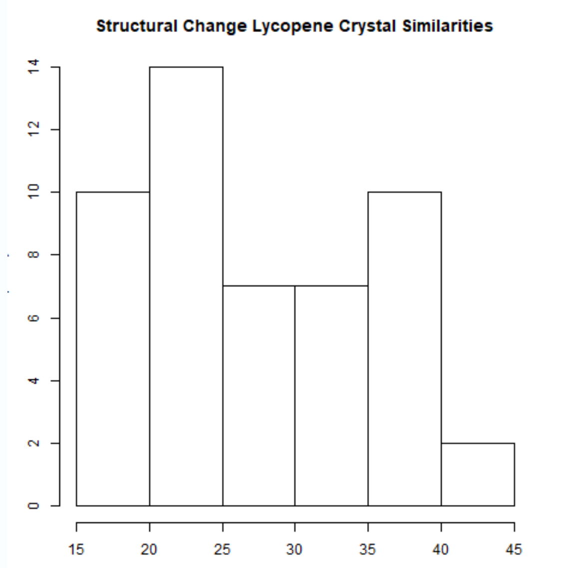

Figure 6: Distribution of the radius of gyration for the overall structural change for 50 Lycopene Crystal Similarities [2][1002]

Figure 5a has the relationship of each of the AA properties for group A with the radius of gyration for examination.

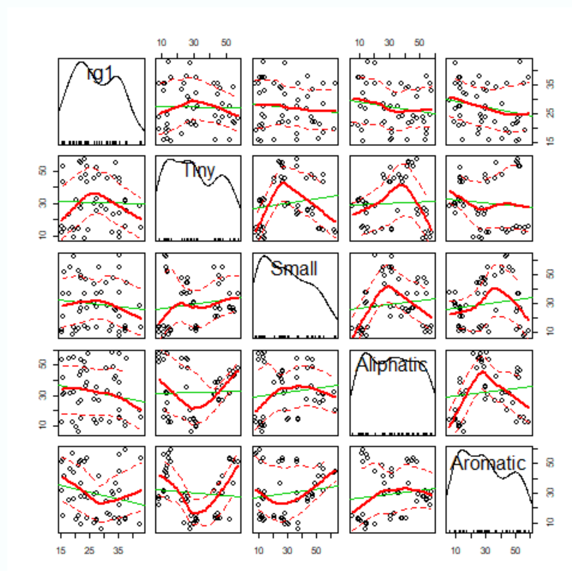

Figure 7: Radius of gyration for the overall structural change for 50 Lycopene Crystal Similarities with Group A [2][1002]

Figure 5b has the relationship of each of the AA properties for group A with the radius of gyration for examination.

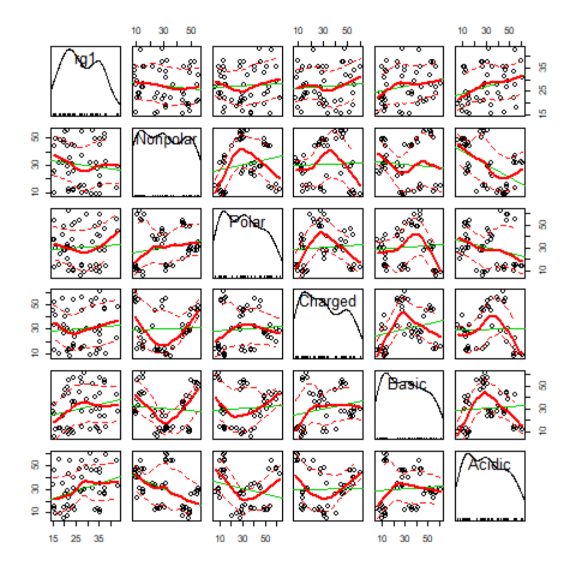

Figure 8: Radius of gyration for the overall structural change for 50 Lycopene Crystal Similarities with Group B [2][1002]

Here a general linear model was used with the ACIDIC property with RG1 with a gaussian distribution with the following results in Table 5.

Deviance Residuals:				
Min	1Q	Median	3Q	Max
-13.559	-5.028	-0.890	6.019	15.433
Coefficients:				
Parameter	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	22.80361	2.23518	10.202	1.31e-13 ***
ACIDIC	0.14592	0.06507	2.243	0.0296 *

Table 2: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1, (Dispersion parameter for gaussian family taken to be 53.75839), Null deviance: 2850.8 on 49 degrees of freedom, Residual deviance: 2580.4 on 48 degrees of freedom, AIC: 345.08. Number of Fisher Scoring iterations: 2 [1000]

Figure 5C has the correlation and boxplots for the radius of gyration and the Model Prediction RG Predict.

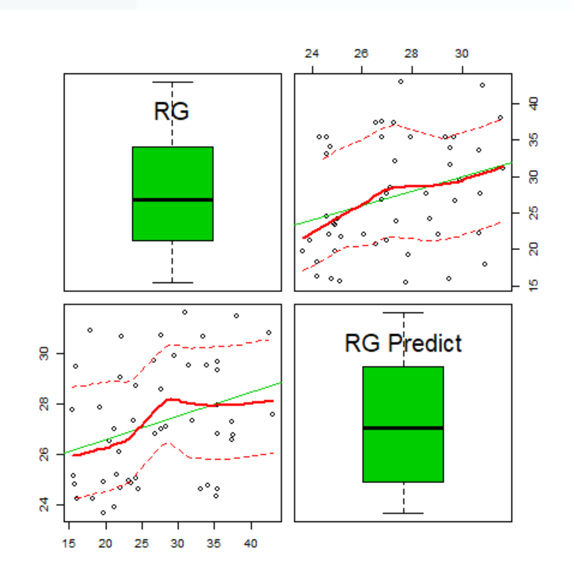

Figure 9: Radius of gyration for the overall structural change for 50 Lycopene Crystal Similarities with GLM Model Prediction [2][1002]

In the design of molecular scales for comparision, consider (a) crucianiProperties [3] (b) kideraFactors [4] (c) zScales [5] (d) FASGAI [6] (e) tScales [7] (f) VHSE [8] (g) protFP [9] (h) stScales [10] (i) BLOSUM [11] and (j) MSWHIM [?key12] [1001]. The Kidera Factors are from multivariate analysis to 188 physical properties of the 20 amino acids with dimensionaltiy reduction techniques. A 10-dimensional vector of orthogonal factors where the first four factors are essentially pure physical properties; the remaining six factors are superpositions of several physical properties are presented in Table 6. [1001]

	Names	HBF	SCS	ESP	H	DBP	PSV	FEP	OAR	PKC	SH
1	1O60	-136	-282	12	59	18	-334	-2	62	22	-49
2	1O61	-180	-132	-35	130	-108	-411	44	128	42	-52
3	1O62	-180	-132	-35	130	-108	-411	44	128	42	-52
4	1O63	-137	-151	162	112	-134	-315	52	4	2	73
5	1O64	-137	-151	162	112	-134	-315	52	4	2	73
6	1O65	-103	-144	-71	84	-164	-282	-1	3	-82	61
7	1O66	-364	-451	-91	-4	-8	-300	49	47	-97	60
8	1O67	-95	-151	-71	73	-180	-277	-10	-2	-83	69
9	1O68	-351	-442	-84	-4	-4	-301	54	60	-95	56
10	1O69	-180	-132	-35	130	-108	-411	44	128	42	-52
11	1O6B	-117	-197	106	125	-132	-294	-77	6	14	-19
12	1O6C	-171	-191	-46	140	-50	-350	-14	13	51	66
13	1O6D	-280	-69	99	146	-113	-318	82	191	9	99
14	1VGT	-209	-335	-66	38	-230	-345	-8	-87	106	112
15	1VGU	-213	-328	-49	55	-250	-323	-6	-105	103	132
16	1VGV	-146	-192	59	77	-129	-281	-0	-64	47	48
17	1VGW	-283	-360	15	107	-229	-409	71	-68	0	-3
18	1VGX	-145	-284	-106	112	-49	-277	6	38	11	140
19	1VGY	-90	-303	-18	108	-143	-268	42	23	-7	11
20	1VGZ	-297	-346	13	115	-231	-409	68	-76	8	-5
21	1VH0	-246	-90	56	214	37	-332	101	34	-110	-121
22	1VH1	-225	-251	68	90	-242	-221	-59	62	2	17
23	1VH2	-157	-273	-99	126	-67	-272	9	34	4	152
24	1VH3	-201	-204	66	102	-108	-355	-90	35	89	-88
25	1VH4	-190	-243	-15	135	-108	-279	-6	9	83	-42
26	1VH5	-170	-294	81	98	-200	-225	71	90	-101	114
27	1VH6	-245	-129	-37	146	-181	-212	102	149	56	-193
28	1VH7	-233	-279	61	118	-202	-381	-3	142	-35	68
29	1VH8	-224	-350	11	72	-188	-272	40	168	67	-28
30	1VH9	-126	-341	68	67	-135	-133	95	12	-94	210
31	1VHA	-224	-348	11	74	-186	-273	41	168	66	-28
32	1VHC	-194	-325	47	-16	-163	-392	18	21	56	-148
33	1VHD	-129	-207	-43	120	-76	-411	25	31	-96	150
34	1VHE	-153	-318	51	120	-66	-257	-36	124	-89	-16
35	1VHF	-319	79	-79	140	-268	-400	133	3	-53	9
36	1VHG	-136	-144	-39	106	-184	-364	-32	-41	-29	79
37	1VHJ	-137	-334	77	71	-76	-215	-54	246	-49	18
38	1VHK	-211	-194	29	134	-183	-271	59	-58	-41	-50
39	1VHL	-269	-303	27	107	-179	-334	-100	-69	76	-112
40	1VHM	-185	-259	57	43	-210	-345	35	99	99	9
41	1VHO	-131	-262	108	86	-85	-375	-36	89	-67	113
42	1VHQ	-290	-422	-39	7	-179	-305	-3	6	26	-7
43	1VHS	-134	-188	-84	58	-263	-323	10	86	-32	50
44	1VHT	-275	-305	21	108	-182	-330	-107	-66	80	-112
45	1VHU	-303	-293	-49	87	-116	-406	61	171	-98	80
46	1VHV	-288	-221	5	105	-170	-382	-45	13	-34	172
47	1VHW	-136	-333	81	79	-74	-209	-52	235	-54	12
48	1VHX	-200	-263	45	116	-125	-393	70	44	-167	45
49	1VHY	-236	-223	23	122	-171	-250	91	-10	-14	-79
50	1VHZ	-134	-145	-45	105	-187	-367	-31	-46	-30	79

Table 3: Values multipled by 1000 with 0 decimal places. HBF=Helix/bend preference, SCS=Side-chain size, ESP=Extended structure preference, H=Hydrophobicity, DBP=Double-bend preference,PSV=Partial specific volume, FEP=Flat extended preference, OAR=Occurrence in alpha region, PKC=pK-C, and SH=Surrounding hydrophobicity. [1001]

A cluster analysis was performed based on the Kidera Factors from Table 6 and the cluster dendrogram presented in Figure 6.

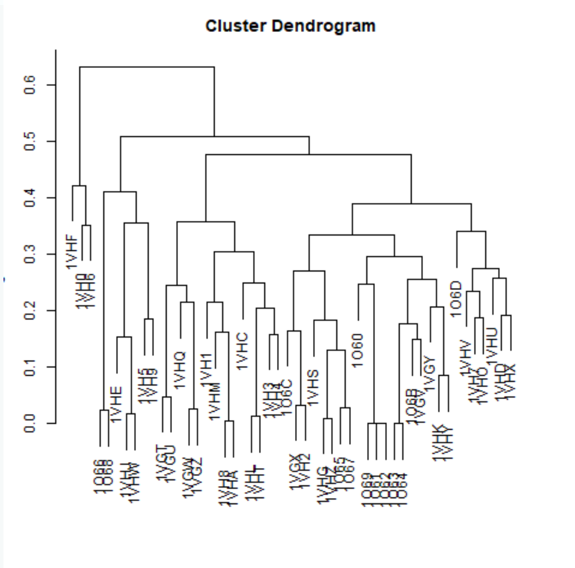

Figure 10: Cluster Analysis of Kidera Factors [1001]

3 Conclusion

A collection of 50 proteins similar to Lycopene 1VHQ crystal were examined with standard molecular properties and compared and contrasted. The modes and fluctuations of the lycopene crystal was examined and a cluster analysis performed for the descriptive statistics with an emphasis on the third moment. In addition, comparisions were made with the overall structural change measured by the radius of gyration for all the crystals.

4 Bibliography

[1] Badger, J et al. “Structural analysis of a set of proteins resulting from a bacterial genomics project.” Proteins vol. 60,4 (2005): 787-96. doi:10.1002/prot.20541

[2] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242.

[3] Cruciani, G., Baroni, M., Carosati, E., Clementi, M., Valigi, R., and Clementi, S. (2004) Peptide studies by means of principal properties of amino acids derived from MIF descriptors. J. Chemom. 18, 146-155.

[4] Kidera, A., Konishi, Y., Oka, M., Ooi, T., and Scheraga, H. A. (1985). Statistical analysis of the physical properties of the 20 naturally occurring amino acids. Journal of Protein Chemistry, 4(1), 23-55.

[5] Sandberg M, Eriksson L, Jonsson J, Sjostrom M, Wold S: New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. J Med Chem 1998, 41:2481-2491.

[6] Liang, G., and Li, Z. (2007). Factor analysis scale of generalized amino acid information as the source of a new set of descriptors for elucidating the structure and activity relationships of cationic antimicrobial peptides. Molecular Informatics, 26(6), 754-763.

[7] Tian F, Zhou P, Li Z: T-scale as a novel vector of topological descriptors for amino acids and its application in QSARs of peptides. J Mol Struct. 2007, 830: 106-115. 10.1016/j.molstruc.2006.07.004.

[8] van Westen, G. J., Swier, R. F., Wegner, J. K., IJzerman, A. P., van Vlijmen, H. W., and Bender, A. (2013). Benchmarking of protein descriptor sets in proteochemometric modeling (part 1): comparative study of 13 amino acid descriptor sets. Journal of cheminformatics, 5(1), 41.

[9] Yang, L., Shu, M., Ma, K., Mei, H., Jiang, Y., and Li, Z. (2010). ST-scale as a novel amino acid descriptor and its application in QSAM of peptides and analogues. Amino acids, 38(3), 805-816.

[10] Georgiev, A. G. (2009). Interpretable numerical descriptors of amino acid space. Journal of Computational Biology, 16(5), 703-723.

[11] Zaliani, A., and Gancia, E. (1999). MS-WHIM scores for amino acids: a new 3D-description for peptide QSAR and QSPR studies. Journal of chemical information and computer sciences, 39(3), 525-533.

[400] Kanehisa, Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K.; KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 45, D353-D361 (2017).

[401] Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M.; KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44, D457-D462 (2016).

[402] Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27-30 (2000).

[403] Petri, V., Jayaraman, P., Tutaj, M., Hayman, G. T., Smith, J. R., De Pons, J., … Jacob, H. J. (2014). The pathway ontology – updates and applications. Journal of Biomedical Semantics, 5, 7. http://doi.org/10.1186/2041-1480-5-7

[601] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, von Mering C. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019 Jan; 47:D607-613.PubMed

[800] Kim, Sunghwan et al. “PubChem in 2021: new data content and improved web interfaces.” Nucleic Acids Res. vol. 49,D1 (2021): D1388-D1395. doi:10.1093/nar/gkaa971

[1000] R Core Team. “R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.” (2015). URL https://www.R-project.org/.

[1001] Osorio, D., Rondon-Villarreal, P. and Torres, R. Peptides: A package for data mining of antimicrobial peptides. The R Journal. 7(1), 4-14 (2015).

[1002] Grant, B.J. et al. (2006) Bioinformatics 22, 2695–2696.