

Figure 1. Pathway of SA for the structure determination of crambin starting with an extended β -strand using interproton distance data.¹³ Simulated interproton distance data were used that could realistically be obtained by two-dimensional NOE spectroscopy. Snapshots of the C^α backbone are shown at 1-ps intervals. During the first 5 ps (thin lines), only interproton distance restraints between residues i , $i \pm 1$, $i \pm 2$, $i \pm 3$, $i \pm 4$, and $i \pm 5$ were taken into account, while all interproton distances were used during the following stages (thick lines).

Mileage Chart

	Atlanta, GA	Atlantic City, NJ	Baltimore, MD	Boston, MA	Charlotte, NC	Chicago, IL	Cincinnati, OH	Cleveland, OH	Columbus, OH	Dallas, TX	Denver, CO	Detroit, MI	Durham, NC	Ft. Worth, TX	Houston, TX	Indianapolis, IN	Kansas City, MO	Las Vegas, NV	Los Angeles, CA	Miami, FL	Milwaukee, WI	Minneapolis, MN	Nashville, TN	New Haven, CT	New Orleans, LA	New York, NY	Orlando, FL	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA	Portland, OR	Raleigh, NC	St. Louis, MO	Salt Lake City, UT	San Antonio, TX	San Diego, CA	San Francisco, CA	San Jose, CA	Seattle, WA	Stamford, CT	Tampa, FL	Washington, DC
Atlanta, GA	0	838	680	1084	250	715	481	727	560	826	1519	741	422	860	875	549	882	2004	2252	682	814	1136	250	967	518	863	449	776	1894	741	2873	410	582	1959	1022	2230	2554	2509	2954	988	458	641
Atlantic City, NJ	838	0	168	370	455	829	640	483	390	1623	1832	670	477	1657	1710	709	1203	2756	2891	1293	911	1241	917	241	1374	165	1069	62	2507	367	2923	480	948	2255	1857	2841	3022	3067	3085	205	1043	202
Baltimore, MD	680	168	0	404	423	696	493	361	428	1460	1624	524	289	1494	1555	574	1070	2473	2724	1146	783	1106	754	277	1206	191	922	106	2332	232	2892	301	800	2120	1702	2736	2899	2944	2932	253	1369	39
Boston, MA	1084	370	404	0	835	976	853	643	769	1868	2008	706	728	1902	1961	948	1442	2810	3130	1547	1069	1399	1168	140	1625	215	1323	308	2746	595	3229	731	1188	2431	2108	3285	3198	3243	3163	376	1427	443
Charlotte, NC	250	455	423	835	0	849	481	728	584	1076	1733	754	120	1110	1114	581	1006	2256	2490	745	961	1283	407	696	736	620	531	524	2091	547	2917	136	740	2149	1307	2480	2801	2756	2873	698	598	384
Chicago, IL	715	829	696	976	849	0	301	343	316	936	1017	279	813	970	1073	182	505	1866	2189	1386	95	416	442	910	938	818	1162	767	1837	468	2250	825	298	1411	1245	2277	2233	2278	2184	874	1215	696
Cincinnati, OH	481	640	493	853	481	301	0	247	103	988	1245	273	531	1022	1115	114	604	1972	2292	1143	391	721	291	709	848	623	940	578	1908	292	2495	543	348	1682	1247	2257	2449	2294	2485	673	972	519
Cleveland, OH	727	483	361	643	728	343	247	0	142	1225	1373	163	590	1259	1358	305	799	2167	2487	1365	439	769	542	565	1132	497	1141	421	2103	128	2599	602	546	1796	1488	2625	2563	2608	2533	529	1208	358
Columbus, OH	560	390	428	769	584	316	103	142	0	1091	1236	180	464	1125	1212	175	608	2256	2358	1236	403	738	945	612	988	555	1022	473	2112	180	2559	529	408	1746	1405	2317	2526	2571	2500	576	966	389
Dallas, TX	826	1623	1460	1868	1076	936	988	1225	1091	0	797	1194	1263	34	241	921	565	1453	1431	1394	1030	940	706	1740	495	1649	1170	1561	1015	1282	2145	1219	843	1240	271	1337	1791	1746	2222	1704	1288	1414
Denver, CO	1519	1832	1624	2008	1733	1017	1245	1373	1236	797	0	1302	1761	763	1038	1047	604	765	1189	2126	1036	871	1205	2033	1292	1852	1902	1770	904	1420	1347	1773	858	497	975	1243	1267	1312	1426	1993	1707	
Detroit, MI	741	670	524	706	794	279	273	163	180	1194	1302	0	677	1228	1326	273	752	2123	2448	1395	363	693	563	606	1143	632	1171	588	2072	294	2523	689	514	1725	1458	2544	2492	2537	2457	570	1224	525
Durham, NC	422	477	289	728	120	813	531	590	464	1263	1761	677	0	1283	1335	660	1312	2370	2648	854	940	1225	510	575	937	500	630	408	2269	469	3007	20	858	2205	1470	2578	2959	2914	2989	539	718	249
Ft. Worth, TX	860	1657	1494	1902	1110	970	1022	1259	1125	34	763	1228	1263	0	236	955	499	1419	1397	1428	1084	934	748	1774	529	1683	1204	1595	981	1316	2111	1297	637	1206	285	1303	1757	1712	2188	1738	1320	1448
Houston, TX	875	1710	1555	1961	1114	1073	1115	1356	1212	241	1038	1326	1335	236	0	1010	714	1484	1564	1306	1150	1181	814	898	363	1742	1042	1648	1206	1396	2368	1347	780	1502	193	1578	1984	1939	2445	862	1045	1501
Indianapolis, IN	549	709	574	948	581	182	114	305	175	921	1047	273	680	955	1010	0	903	1910	2182	1219	281	598	283	744	831	729	995	647	1789	356	2385	680	242	1528	1192	2151	2353	2398	2375	708	949	572
Kansas City, MO	882	1203	1070	1442	1006	505	604	799	668	505	604	752	1312	499	714	503	0	1417	1631	1485	585	435	584	1299	809	1223	1278	1141	1274	858	1993	1132	255	1105	788	1651	1903	1858	1984	1263	1292	1066
Las Vegas, NV	2004	2756	2473	2810	2250	1866	1972	2167	2258	1453	765	2123	2370	1419	1494	1910	1417	0	288	2582	2038	1688	1843	2667	1740	2591	2328	2509	288	2224	1027	2386	1688	418	1283	330	576	531	1384	2631	2434	
Los Angeles, CA	2252	2891	2742	3130	2498	2189	2292	2487	2356	1431	1189	2448	2648	1397	1564	2182	1631	288	0	2885	2149	2033	2098	2987	1947	2911	2661	2289	379	2944	1018	2868	1942	706	1388	126	377	332	1193	2951	2554	2754
Miami, FL	682	1293	1146	1547	745	1386	1143	1365	1238	1394	2126	1395	854	1428	1308	1219	1485	2582	2885	0	1478	1786	921	1401	881	1325	224	1231	2417	1261	3438	834	1231	2621	1435	2881	3238	3193	3469	1365	278	1096
Milwaukee, WI	814	911	783	1069	961	95	391	439	403	1030	1036	383	940	1064	1150	281	585	2038	2149	1478	0	335	554	1000	1018	931	1252	849	1846	333	2138	928	370	1434	1307	2278	2186	2231	2043	864	1206	786
Minneapolis, MN	1138	1241	1106	1399	1283	416	721	769	736	940	871	693	1225	934	1181	598	435	1688	2033	1786	335	0	876	1337	1228	1261	1562	1179	1689	885	1838	424	953	1309	1388	2083	2077	2122	1691	1301	1615	1116
Nashville, TN	250	917	754	1168	407	442	291	542	945	706	1205	553	510	740	814	283	584	1843	2098	921	554	876	0	1025	549	949	697	855	1782	582	2517	338	310	1700	937	2043	2409	2364	2548	989	664	708
New Haven, CT	967	241	277	140	696	910	709	565	612	1740	2033	606	575	1774	898	744	1289	2667	2987	1401	1000	1337	1025	0	1482	76	723	184	2603	443	3184	587	1888	2351	1985	2927	3158	3293	3181	35	1893	326
New Orleans, LA	518	1374	1206	1625	736	938	846	1132	988	495	1292	1143	937	529	383	831	809	1740	1947	881	1018	1228	549	1482	0	1406	657	1312	1853	1172	2694	987	688	1738	874	1834	2388	2393	2791	1446	639	1165
New York, NY	863	165	191	215	620	818	623	497	555	1649	1852	632	500	1683	1742	729	2223	2591	2911	1325	931	1261	949	76	1406	0	1101	91	2527	376	3088	512	968	2275	1889	2451	3082	3127	3028	607	1208	236
Orlando, FL	449	1069	922	1323	531	1182	940	1141	1022	1170	1902	1171	630	1204	1042	995	1278	2328	2661	224	1252	1562	697	723	657	1101	0	1007	2193	1837	3214	818	1007	2397	1211	2857	3014	2688	3245	697	78	872
Philadelphia, PA	776	62	108	308	524	787	578	421	473	1561	1770	588	408	1595	1848	647	1141	2509	2829	1231	849	1179	855	184	1312	91	1007	0	2446	305	3088	418	887	2193	1795	2779	2960	3005	2943	148	1114	840
Phoenix, AZ	1894	2507	2332	2746	2091	1837	1908	2103	2112	1015	904	2072	2269	981	1208	1769	1274	288	379	2417	1945	1669	1782	2603	1553	2527	2193	2445	0	2188	1383											

pg. 4 & 5
Alaska



pg. 6 & 7
Puerto Rico,
U.S. Virgin Islands
and Hawaii

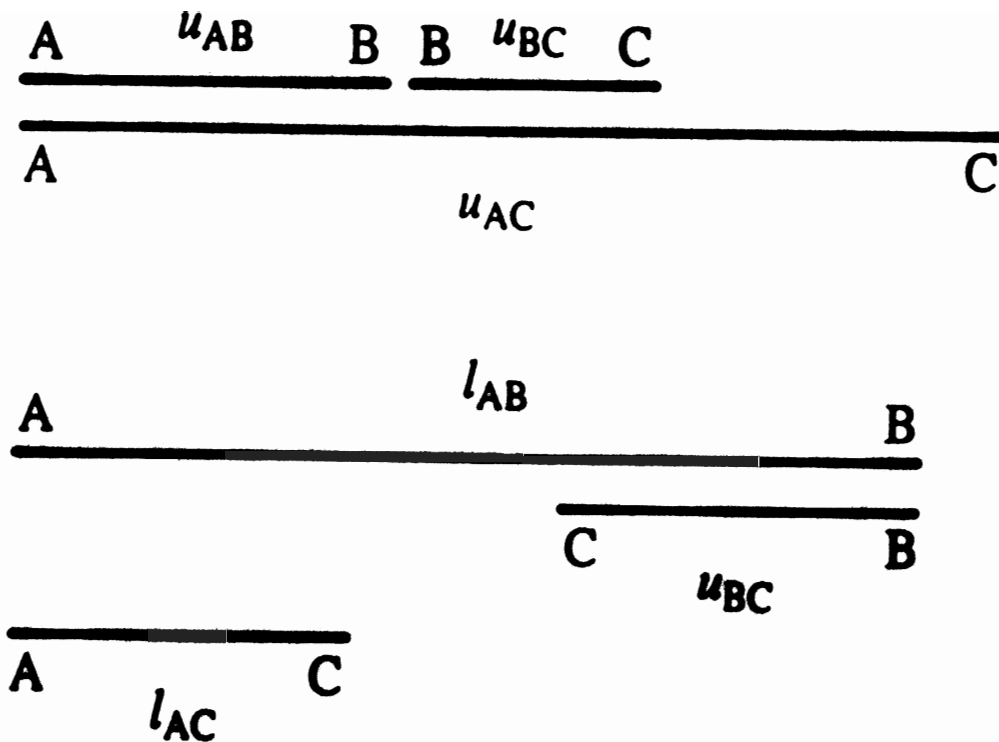


Fig. 8.14 The two triangle inequalities used in distance geometry.

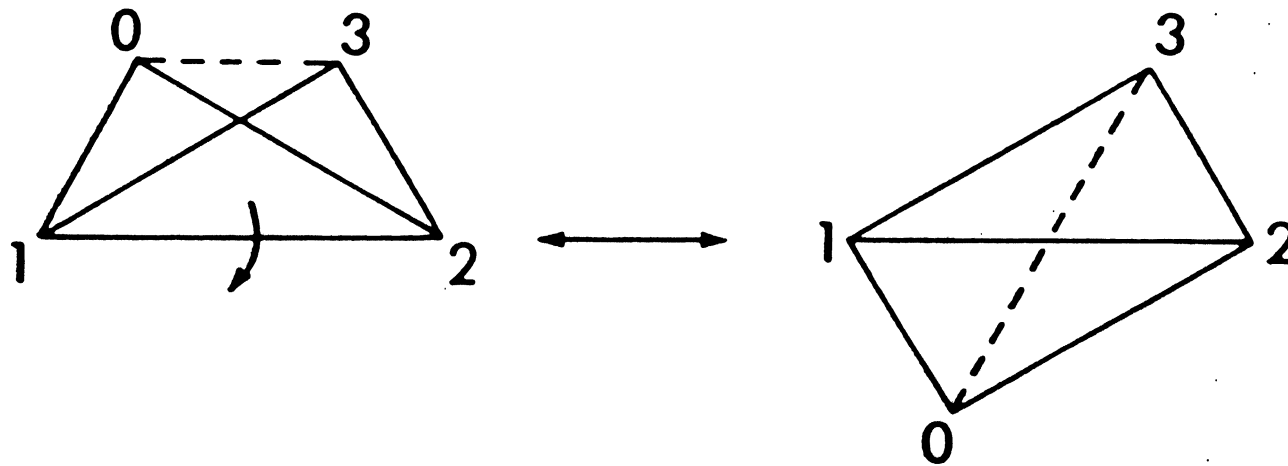
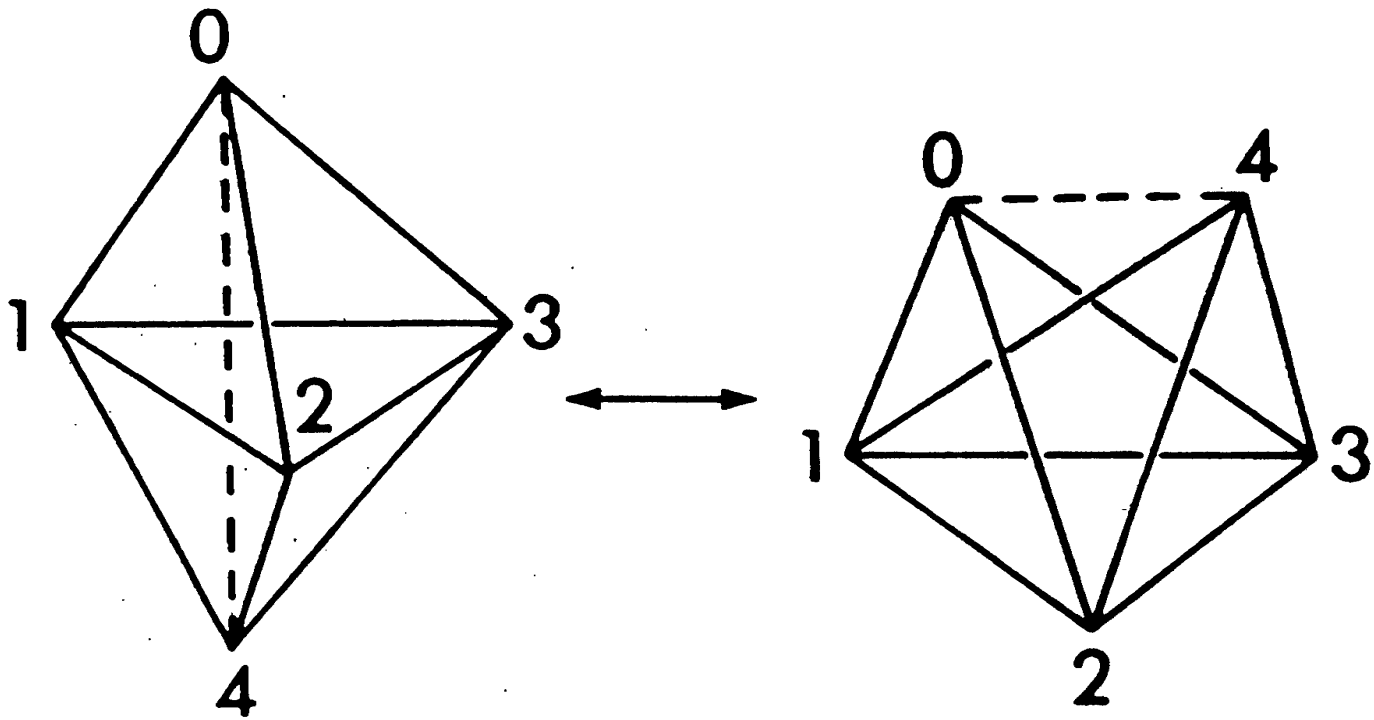


Figure 1. The two possible solutions of the tetrangle equality for a configuration of four points.



The two possible solutions of the pentangle equality for a configuration of five points.

DISTANCE GEOMETRY PROTOCOL

INPUT

Distance, Chirality and Torsion Constraints

BOUNDS SMOOTHING

Upper Triangle, Lower Triangle, Quadrangle

TRIAL DISTANCE SELECTION

Choose Distances between the Bounds

EMBEDDING

Compute Centroid, Metric Matrix, Principal Components, then Atomic Coordinates

REFINEMENT

Nonlinear Optimization to satisfy Constraints

KEY EQUATIONS

Metric Matrix = $g_{ij} = x_i \cdot x_j$

$$d_{i0}^2 = 1/N \sum d_{ij}^2 - 1/N^2 \sum d_{jk}^2$$

$$g_{ij} = (d_{i0}^2 + d_{j0}^2 - d_{ij}^2) / 2$$

$$g_{ij} = \sum x_{ik} x_{jk} = \sum w_{ik} w_{jk} \lambda_k$$

$$x_{ik} = \lambda_k^{1/2} \cdot w_{ik}$$

METRIZATION METHODS

**Random Fractional
Triangle Correlated
Ordered Atom Based
Four-Point Atom-Based
Random Pair-Based (*)**

**(*) Uses a Modified Murchland
Shortest Path Update Algorithm**

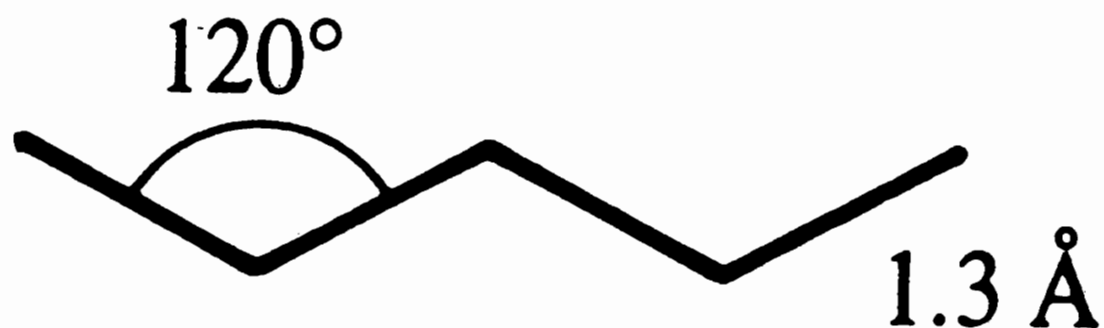


Fig. 8.15 Five-carbon fragment to illustrate distance geometry

initial bounds matrix

$$\begin{pmatrix} 0.0 & 1.3 & 2.2517 & 3.4395 & 99.0 \\ 1.3 & 0.0 & 1.3 & 2.2517 & 3.4395 \\ 2.2517 & 1.3 & 0.0 & 1.3 & 2.2517 \\ 2.6 & 2.2517 & 1.3 & 0.0 & 1.3 \\ 2.8 & 2.6 & 2.2517 & 1.3 & 0.0 \end{pmatrix}$$

smoothed bounds matrix

$$\begin{pmatrix} 0.0 & 1.3 & 2.2517 & 3.4395 & 4.5033 \\ 1.3 & 0.0 & 1.3 & 2.2517 & 3.4395 \\ 2.2517 & 1.3 & 0.0 & 1.3 & 2.2517 \\ 2.6 & 2.2517 & 1.3 & 0.0 & 1.3 \\ 2.8 & 2.6 & 2.2517 & 1.3 & 0.0 \end{pmatrix}$$

randomly assigned distance matrix

$$\begin{pmatrix} 0.0 & 1.3 & 2.25 & 3.11 & 3.42 \\ & 0.0 & 1.3 & 2.25 & 2.85 \\ & & 0.0 & 1.3 & 2.25 \\ & & & 0.0 & 1.3 \\ & & & & 0.0 \end{pmatrix}$$

We calculate the *metric matrix*, G , each of whose elements (i, j) is equal to the scalar product of the vectors from the origin to atoms i and j :

$$G_{ij} = \mathbf{i} \cdot \mathbf{j} \quad (8.4)$$

The elements G_{ij} can be calculated from the distance matrix using the cosine rule:

$$G_{ij} = (d_{i0}^2 + d_{j0}^2 - d_{ij}^2)/2 \quad (8.5)$$

d_{i0} is the distance from the origin to atom i and d_{ij} is the distance between atoms i and j .

It is usual to take the centre of the molecule as the origin of the coordinate system. The distance of each atom from the centre can be calculated directly from the interatomic distances using the following expression:

$$d_{i0}^2 = \frac{1}{N} \sum_{j=1}^N d_{ij}^2 - \frac{1}{N^2} \sum_{j=2}^N \sum_{k=1}^{j-1} d_{jk}^2 \quad (8.6)$$

The metric matrix G is a square symmetric matrix. A general property of such matrices is that they can be decomposed as follows:

$$G = \mathbf{V} \mathbf{L}^2 \mathbf{V}^T \quad (8.7)$$

The diagonal elements of \mathbf{L}^2 are the eigenvalues of G and the columns of \mathbf{V} are its eigenvectors. The atomic coordinates can be derived from the metric matrix by rewriting equation (8.4) as

$$G = \mathbf{X} \mathbf{X}^T \quad (8.8)$$

\mathbf{X} is a matrix containing the atomic coordinates. Equating equations (8.7) and (8.8) gives $\mathbf{X} = \mathbf{V} \mathbf{L}$

$$\mathbf{X} = \mathbf{V} \mathbf{L} \quad (8.9)$$

As \mathbf{L} has only diagonal entries, the matrix \mathbf{L} is identical to its transpose: $\mathbf{L} = \mathbf{L}^T$. The atomic coordinates are thus obtained by multiplying the square roots of the eigenvalues by the eigenvectors.

metric matrix

$$\begin{pmatrix} 3.571 & 1.569 & -0.427 & -2.276 & -2.436 \\ 1.569 & 1.256 & 0.105 & -1.122 & -1.808 \\ -0.427 & 0.105 & 0.644 & 0.261 & -0.583 \\ -2.276 & -1.122 & 0.261 & 1.569 & 1.569 \\ -2.436 & -1.808 & -0.583 & 1.569 & 3.259 \end{pmatrix}$$

eigenvalues

8.18, 1.74, 0.26, 0.10 and 0.0

matrix of eigenvectors

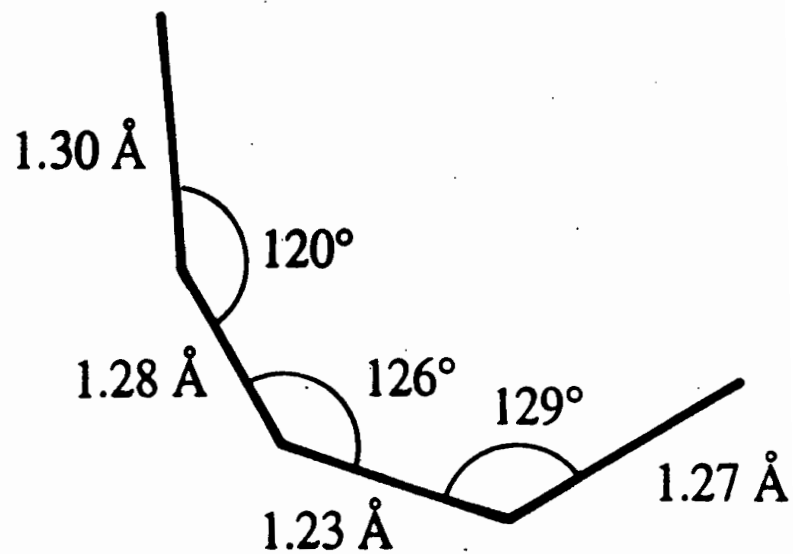
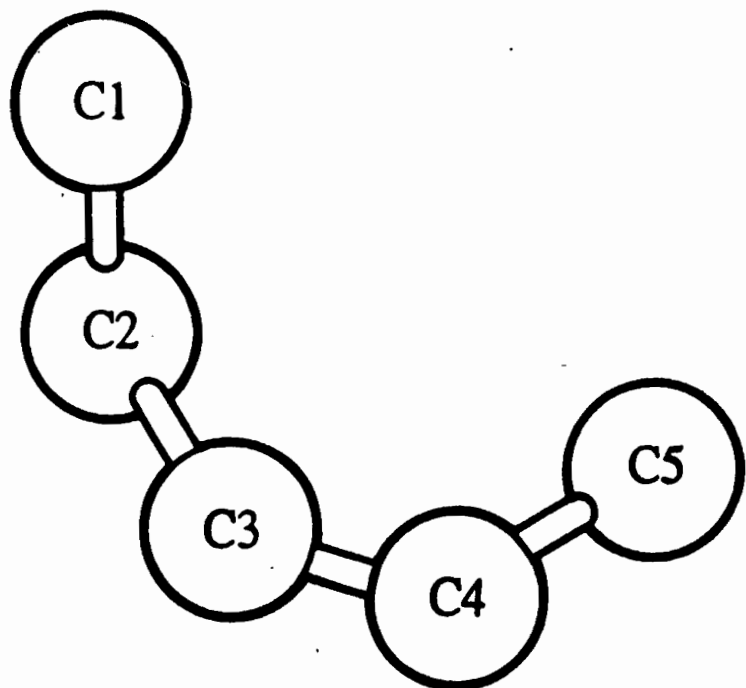
$$\mathbf{W} = \begin{pmatrix} 0.621 & 0.455 & -0.425 & 0.164 \\ 0.355 & -0.184 & 0.800 & 0.020 \\ 0.0 & -0.573 & -0.368 & -0.580 \\ -0.408 & -0.287 & -0.153 & 0.727 \\ -0.567 & 0.590 & 0.145 & -0.330 \end{pmatrix}$$

$$x_i = \sqrt{\lambda_1} W_{i1}$$

$$y_i = \sqrt{\lambda_2} W_{i2}$$

$$z_i = \sqrt{\lambda_3} W_{i3}$$

<i>Atom</i>	<i>x coordinate</i>	<i>y coordinate</i>	<i>z coordinate</i>
1	1.777	0.601	- 0.218
2	1.014	- 0.244	0.410
3	- 0.001	- 0.757	- 0.188
4	- 1.166	- 0.379	- 0.079
5	- 1.623	0.799	0.075



$$\begin{pmatrix}
 0.0 & 1.299 & 2.24 & 3.10 & 3.42 \\
 & 0.0 & 1.29 & 2.24 & 2.85 \\
 & & 0.0 & 1.23 & 2.25 \\
 & & & 0.0 & 1.25 \\
 & & & & 0.0
 \end{pmatrix}$$

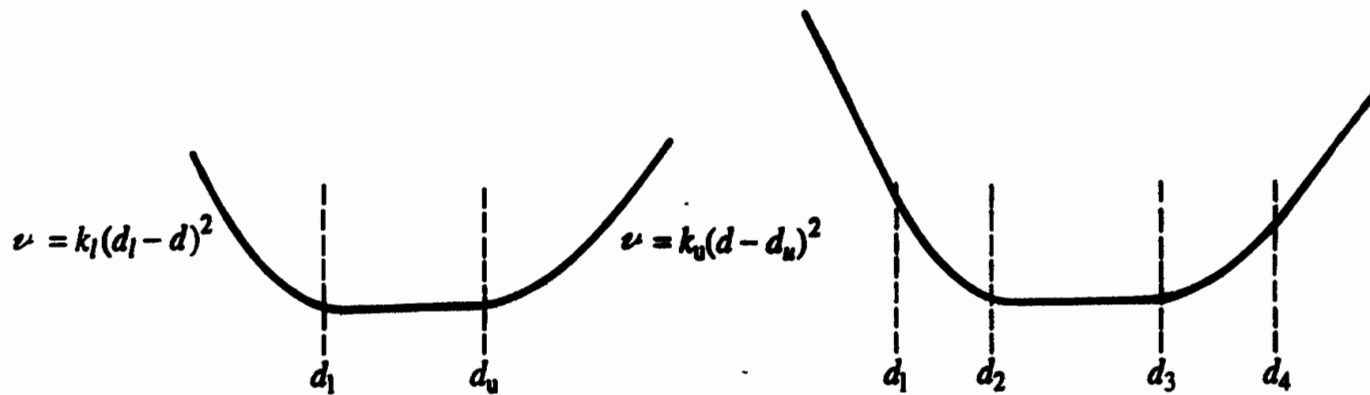


Fig. 8.23 A restraining potential that does not penalise structures in which the distance lies between the lower and upper distances d_l and d_u and uses harmonic functions outside this range (left). The harmonic potentials may also be replaced by linear restraints further from this region (right).

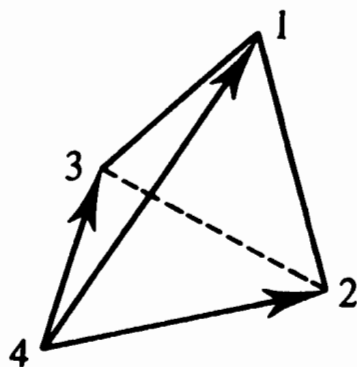


Fig. 8.17 The stereochemistry about tetrahedral atoms can be maintained with an appropriate chiral constraint.

During the optimisation of the structure against the distance constraints it is usual to incorporate *chiral constraints*. These are used to ensure that the final conformation is the desired stereoisomer. Chiral constraints are necessary because the interatomic distances in two enantiomeric conformations are identical and as a consequence the 'wrong' isomer may quite legitimately be generated. Chiral constraints are usually incorporated into the error function as a chiral volume, calculated as a scalar triple product. For example, to maintain the correct stereochemistry about the tetrahedral atom number 4 in Figure 8.17, the following scalar triple product must be positive:

$$(\mathbf{v}_1 - \mathbf{v}_4) \cdot [(\mathbf{v}_2 - \mathbf{v}_4) \times (\mathbf{v}_3 - \mathbf{v}_4)] \quad (8.21)$$

The other stereoisomer corresponds to a negative chiral volume. Chiral constraints are included in the penalty function by adding terms of the following form:

$$(V_{\text{ch}} - V_{\text{ch}}^*)^2 \quad (8.22)$$

V_{ch}^* is the desired value of the chiral constraint. Chiral constraints can also be used to force groups of atoms to lie in the same plane by requiring the chiral volume to have a value of zero.

a



b

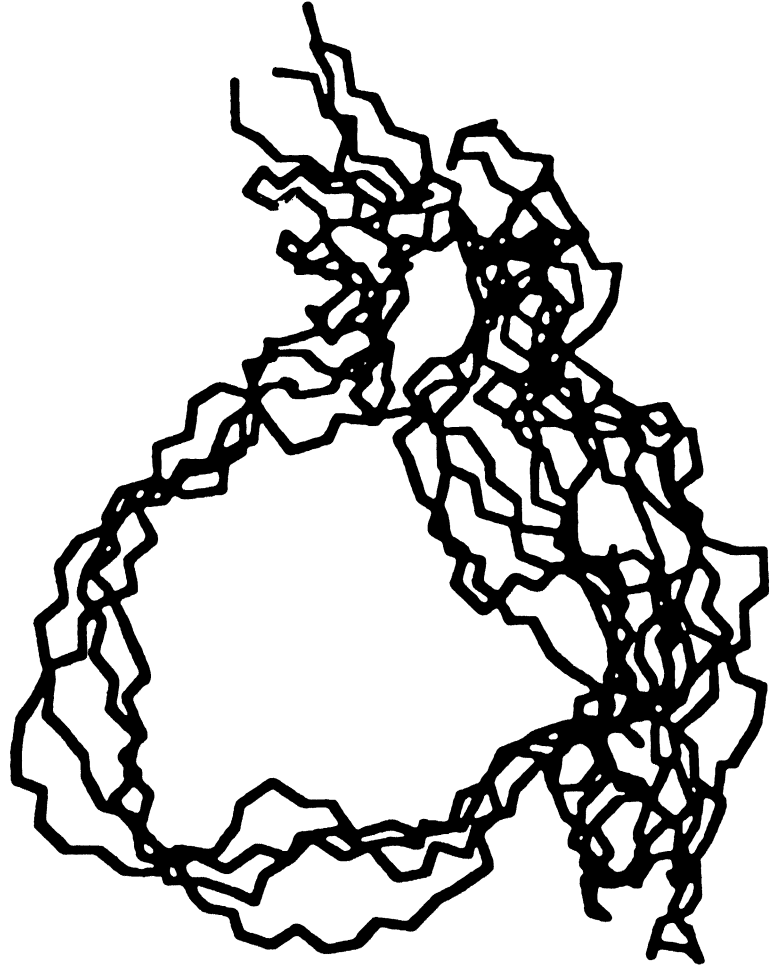
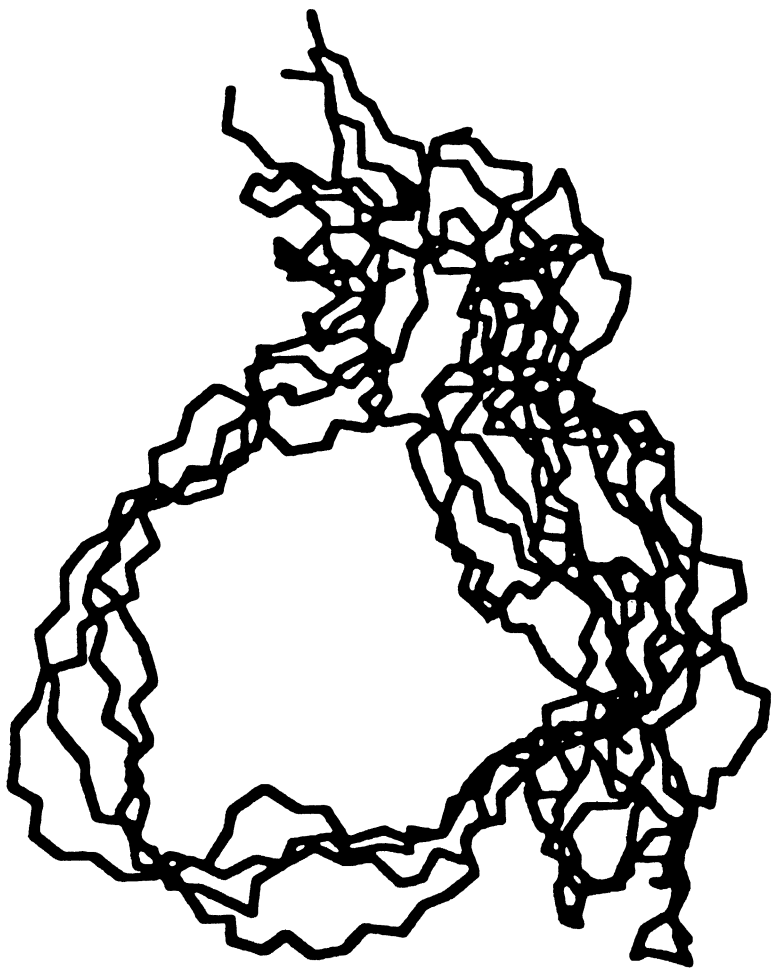


c



d





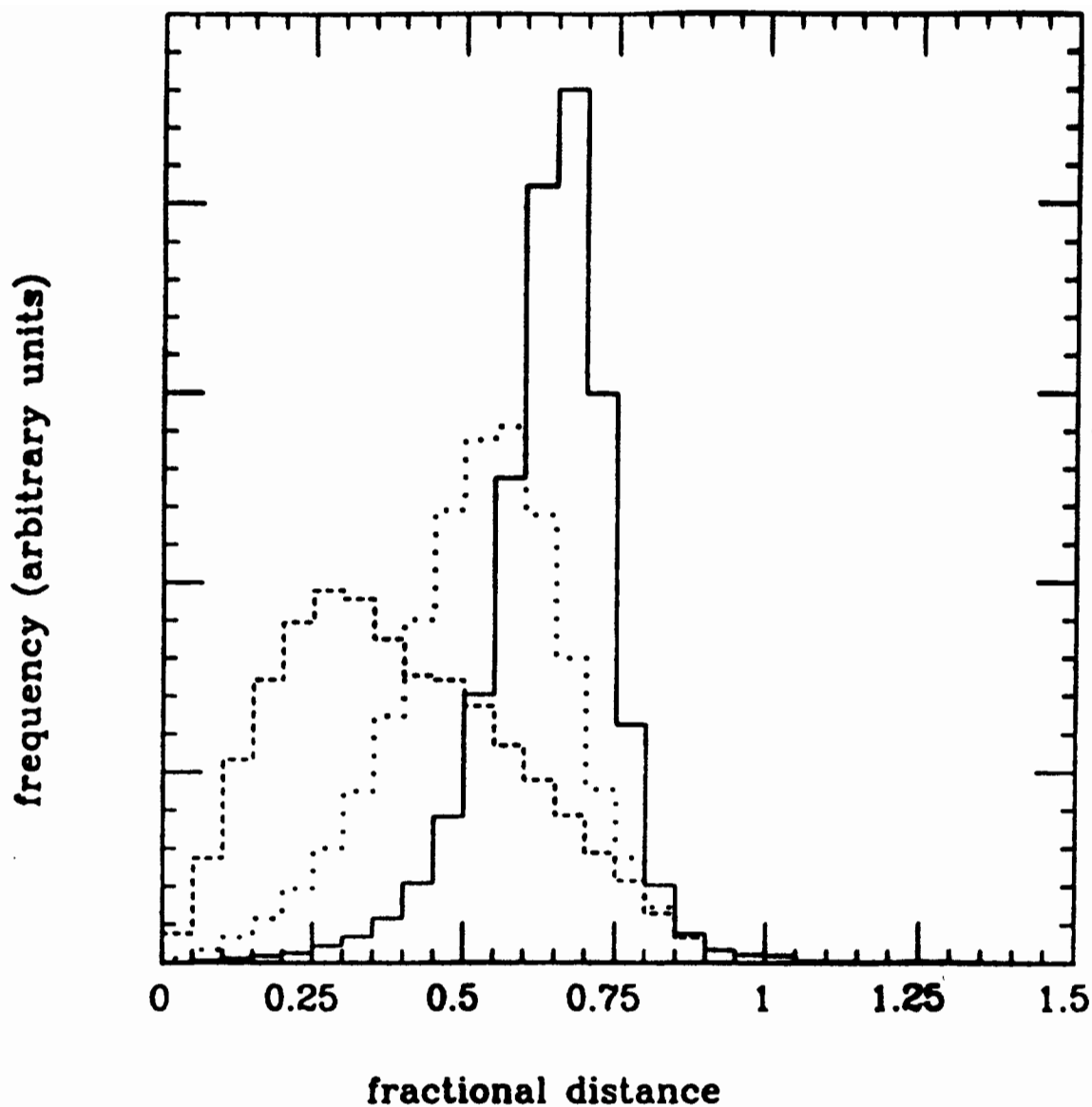


Figure 1. Histogram for the fractional distance, $(\text{distance}(\text{crystal}) - \text{lower bound}) / (\text{upper bound} - \text{lower bound})$. Bounds taken from the all-(H-H) data set (solid line), 10% data set (dotted line) or disulfide-only data set (dashed line).

**FOUR-ATOM METRIZATION
UNIFORM TRIAL DISTRIBUTION**

**PAIRWISE-5% METRIZATION
GAUSSIAN TRIAL DISTRIBUTION**

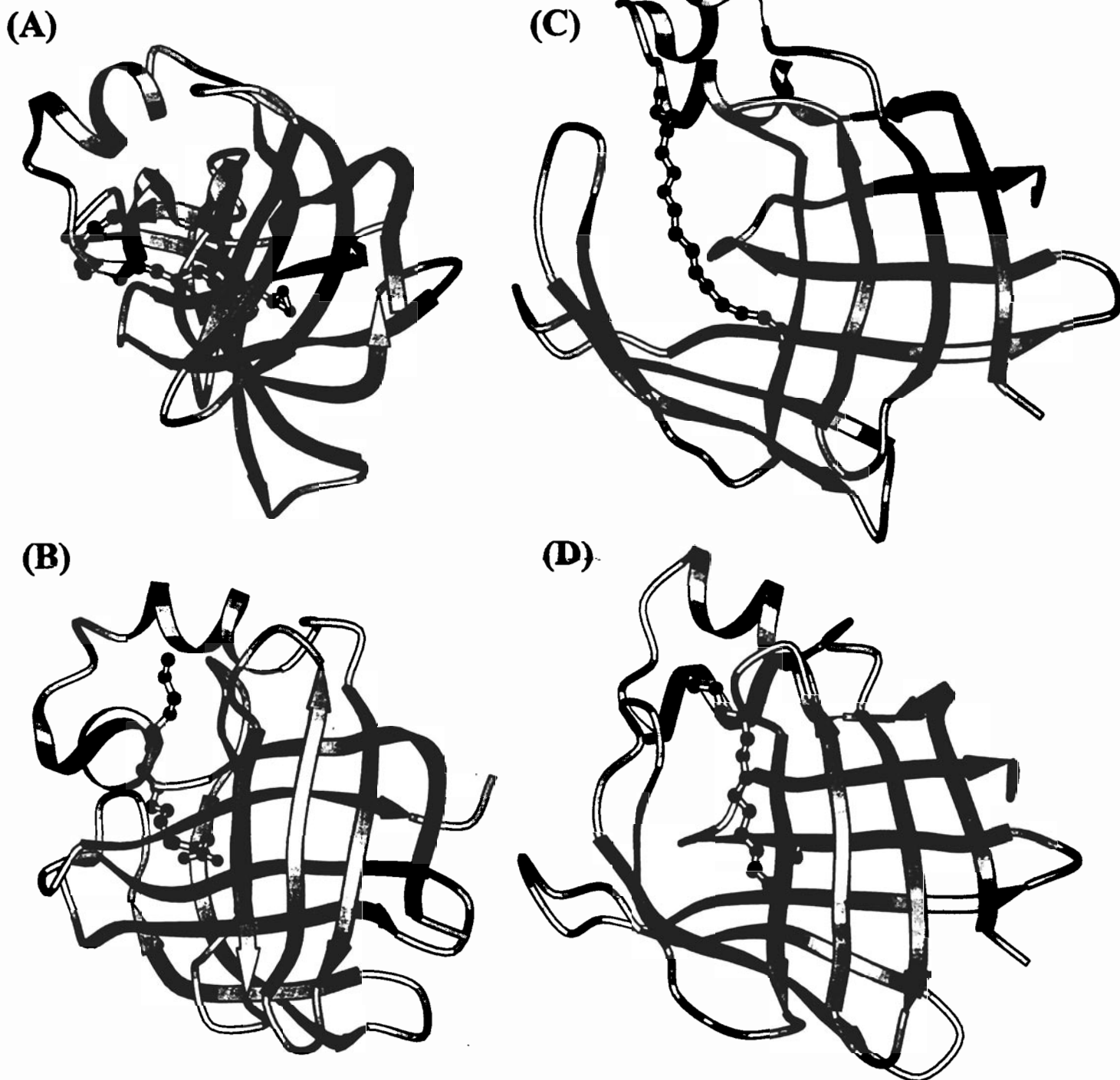


Figure 10. Ribbon diagrams illustrating the influence of the distance geometry metrization algorithm on the calculated structures both prior to ((A) and (C)) and after ((B) and (D)) simulated annealing refinement. All of the calculations were performed using DISTGEOM, which has the option of altering the trial distribution and the type of metrization employed. The structures in (A) and (B) were calculated using four-atom metrization with a uniform trial distribution; this is the type of algorithm implemented in X-PLOR version 3.1. The structures in (C) and (D) were calculated using pairwise-5% metrization with an optimized Gaussian trial distribution. This latter algorithm was used in the current study. For reference, the correct positions of secondary structure elements, as defined by $^1\text{H}/^{13}\text{C}$ chemical shifts, are indicated in each ribbon diagram; this is not meant to imply that regions of the calculated structures shown here actually adopt ϕ/ψ angles consistent with regular α -helix and β -sheet, especially in the distorted structures shown on the left. All of the calculations used the final set of distance restraints listed in Table 3. The structures shown are examples from an ensemble of structures obtained for each set of calculations; other members of the ensemble showed the same general characteristics as shown here.

Protein Engineering: Map Distance Geometry Problem

Initial Table of Intercity Mileages :

	A	B	C	D	F	H	M	P
A	0	1084	715	1519	860	875	814	776
B	1084	0	976	2008	1902	1961	1069	308
C	715	976	0	1017	970	1073	95	767
D	1519	2008	1017	0	763	1038	1036	1770
F	860	1902	970	763	0	236	1064	1595
H	875	1961	1073	1038	236	0	1150	1648
M	814	1069	95	1036	1064	1150	0	849
P	776	308	767	1770	1595	1648	849	0

Applying Triangle Inequality to the Distances :

A to M Distance Changed from 814 to 810 via A --> C --> M
 B to H Distance Changed from 1961 to 1959 via B --> A --> H
 B to H Distance Changed from 1959 to 1956 via B --> P --> H
 B to D Distance Changed from 2008 to 1993 via B --> C --> D
 D to H Distance Changed from 1038 to 999 via D --> F --> H

Distance Geometry Metric Matrix :

	A	B	C	D	F	H	M	P
A	245203	159156	-108504	-512045	53837	117256	-126632	171728
B	159156	1248166	172302	-842907	-883683	-911418	131519	926866
C	-108504	172302	49014	26397	-144907	-173690	98811	80578
D	-512045	-842907	26397	1038068	528985	397501	61202	-697201
F	53837	-883683	-144907	528985	602071	650655	-186196	-620762
H	117256	-911418	-173690	397501	650655	754934	-204967	-630270
M	-126632	131519	98811	61202	-186196	-204967	157632	68631
P	171728	926866	80578	-697201	-620762	-630270	68631	700430

Distance from each City to the Map Centroid :

City	Distance to Centroid
A	495
B	1117
C	221
D	1019
F	776
H	869
M	397
P	837

Eigenvalues and Eigenvectors of Metric Matrix :

-164207.0202	-31467.5657	0.0000	4993.2342
0.7015	-0.1262	0.3536	-0.2326
-0.2545	-0.4255	0.3536	-0.2421
0.0964	0.7715	0.3536	-0.2294
0.3626	-0.2271	0.3536	0.1407
-0.4811	0.0921	0.3536	-0.4662
-0.2226	-0.1110	0.3536	0.4979
-0.1213	-0.2468	0.3536	-0.0569
-0.0811	0.2730	0.3536	0.5885
14258.5892	96761.3462	1028672.7147	3846507.8266
0.1134	0.0399	-0.5400	0.0821
-0.4188	-0.2519	0.0845	0.5697
-0.3758	0.1587	0.2177	0.0641
-0.0008	-0.3632	0.6044	-0.4182
0.3685	-0.2895	-0.2173	-0.3885
-0.4703	0.1887	-0.3910	-0.3944
0.3353	0.7708	0.2929	0.0617
0.4485	-0.2535	-0.0513	0.4236

Distance Geometry Generated Map Coordinates :

City	X-Coordinate	Y-Coordinate
A	161	-548
B	1117	86
C	126	221
D	-820	613
F	-762	-220
H	-774	-397
M	121	297
P	831	-52

Final Table of Intercity Mileages :

	A	B	C	D	F	H	M	P
A	0	1147	769	1520	979	947	846	833
B	1147	0	1001	2008	1904	1951	1018	318
C	769	1001	0	1024	991	1091	76	756
D	1520	2008	1024	0	835	1011	993	1780
F	979	1904	991	835	0	177	1023	1602
H	947	1951	1091	1011	177	0	1132	1641
M	846	1018	76	993	1023	1132	0	791
P	833	318	756	1780	1602	1641	791	0

Map Showing the Locations of the Cities :

