

3. METHODOLOGY

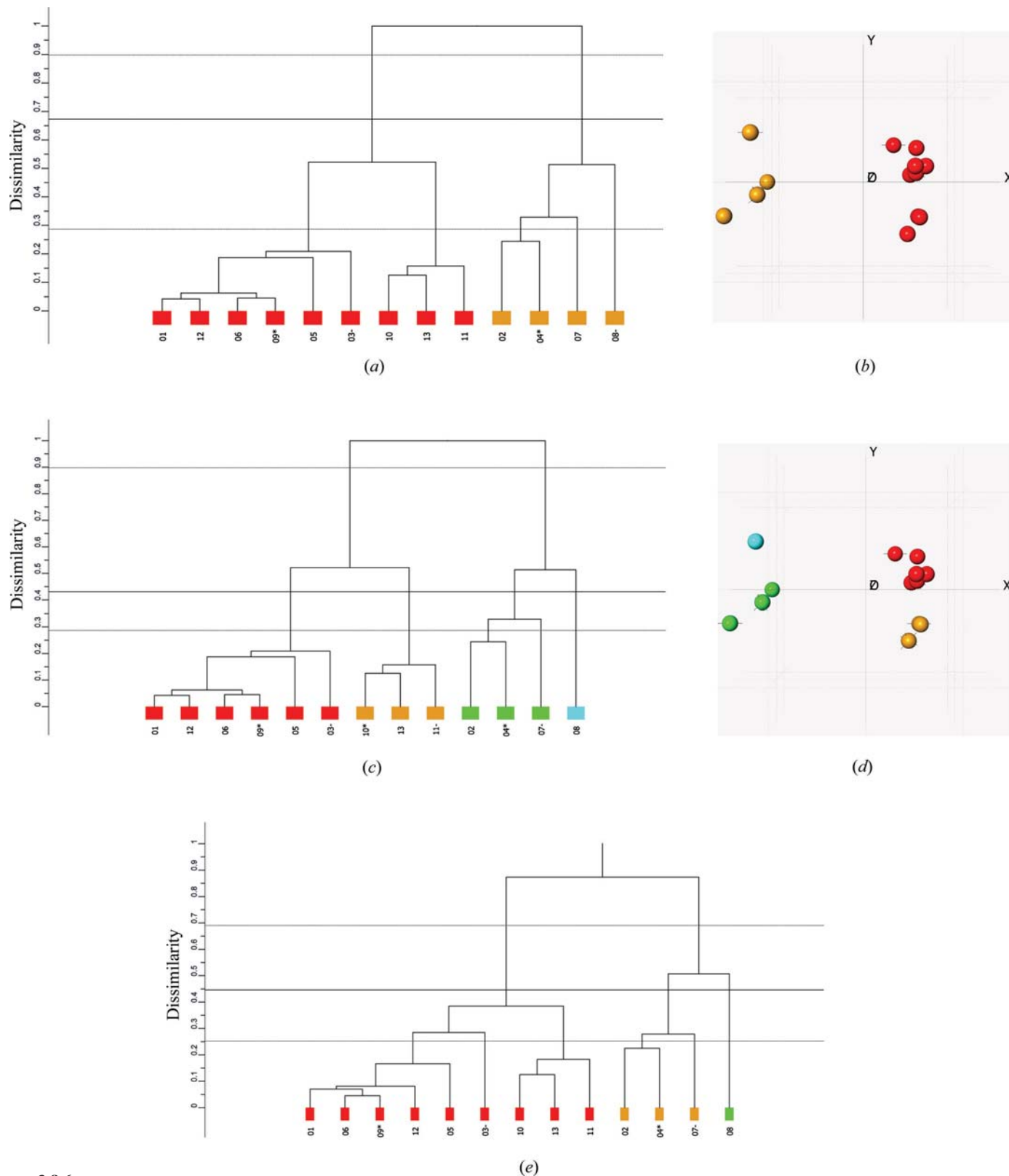


Figure 3.8.6

(a) The initial default dendrogram using the centroid clustering method on 13 PXRD patterns from 13 commercial aspirin samples. (b) The corresponding MDS plot. It can be seen that both clusters have a natural break in them and should be partitioned into two clusters. (c) The dendrogram cut line is reduced. (d) The corresponding MDS plot. The red cluster is now partitioned into two; the remaining patterns are a light-blue singleton and a green triplet cluster. (e) The default dendrogram using the single-link method.

ficient of cluster membership. To compute them, the dissimilarity matrix, δ , is used. If the pattern i belongs to cluster C_r which contains n_r patterns, we define

$$a_i = \sum_{\substack{j \in C_r \\ j \neq i}} \delta_{ij} / (n_r - 1). \tag{3.8.19}$$

This defines the average dissimilarity of pattern i to all the other patterns in cluster C_r . Further define

$$b_i = \min_{s \neq r} \left\{ \sum_{j \in C_s} \delta_{ij} / n_s \right\}. \tag{3.8.20}$$

The silhouette for pattern i is then