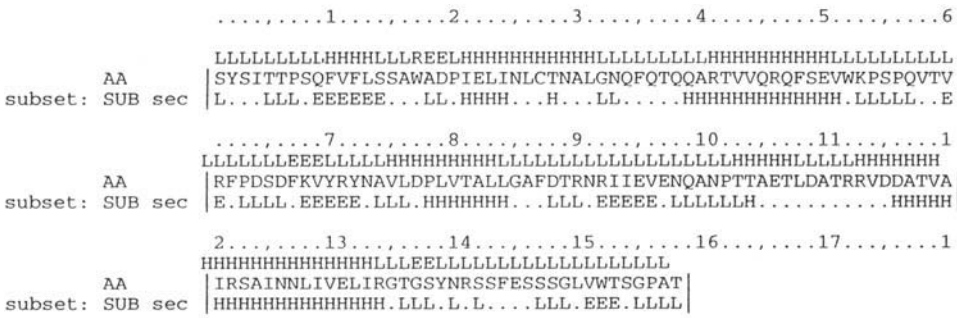

Fig. 3. Protein secondary structure prediction. The predicted secondary structure (below the sequence) and experimentally found secondary structure elements (above the sequence) are shown for the TMV capsid protein. H, helix; L, loop; and E, extended conformation (β-sheet).

tantly related plant virus capsid proteins mostly fall within the boundaries of the high-complexity segments. Low-complexity regions generally correspond to nonglobular protein domains. In practical terms, delineation of low-complexity regions is of major importance for sequence similarity searches, because these regions very frequently produce artifactual significant alignments in database searches ***(3,10,12)***.

2.1.3. Secondary Structure Prediction and Hydropathy Analysis

In spite of very significant effort and availability of several popular methods ***(13,14)***, the problem of secondary structure prediction on the basis of protein sequence is not solved. The recently developed neural network method, which incorporates information from multiple protein sequence alignments, signified marked progress, with the accuracy of prediction apparently reaching 70–80%, depending on the number of available homologous sequences ***(15–17)***. **Figure 3** shows the secondary structure predicted for the tobacco mosaic virus (TMV) coat protein (CP) produced using the PHD program ***(15)***, and compared to the experimentally determined structure. Clearly, the program predicts most (but not all) of the secondary structure elements, even though their boundaries cannot be predicted precisely.

Hydropathy analysis helps to identify portions of the polypeptide chain that are likely to be buried in the protein globule or exposed on its surface, and to detect potential transmembrane segments. Identification of exposed regions in proteins is particularly important for predicting their antigenic properties. The classical methods for hydropathy analysis ***(18,19)*** plot the hydropathy along the sequence, using a sliding window and amino acid hydropathy tables. The