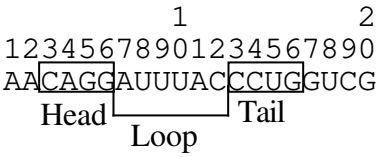



**Figure 1.** Description of a palindrome ((3,16)(4,15)(5,14)(6,13)) located in a sequence of bases (numbers are positions in the sequence). The head of the palindrome corresponds to positions (3,4,5,6), whereas the tail and the loop respectively correspond to positions (13,14,15,16) and (7,8,9,10,11,12).

---

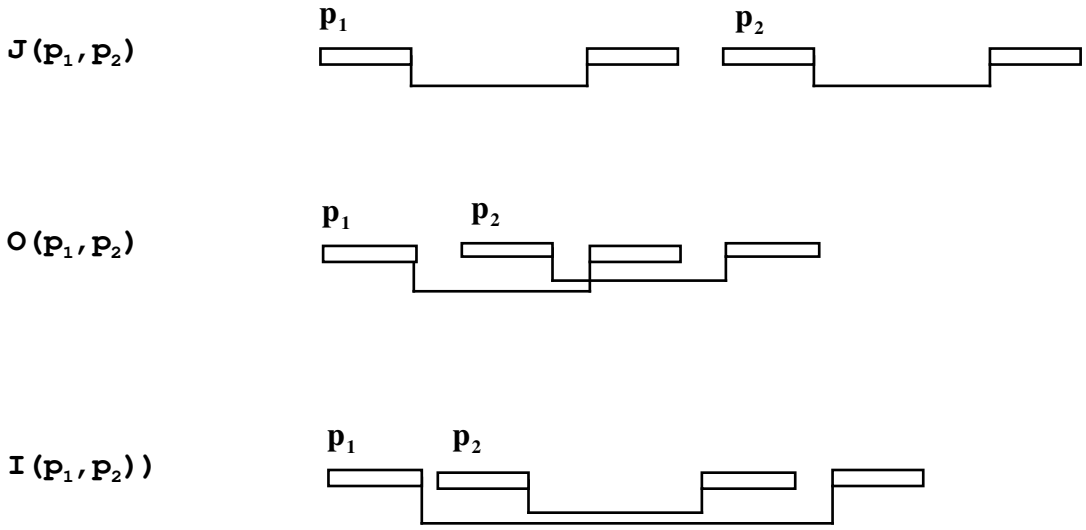



**Figure 2.** Possible structural relations between two palindromes.

---

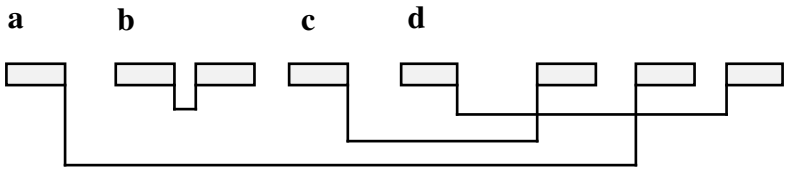



**Figure 3.**Example of a 4-length structural pattern **t**:<**abcd**>.

---

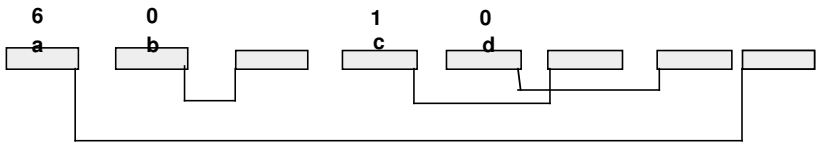



**Figure 4.** Encoding of a structure. Consider the encoding of the structure **t**:<**abcd**> of the Figure 3: each number (standing for a partial structural pattern) is above the name of the palindrome it is associated with : the tail of **a** is preceded by **6** elements (heads or tails)

belonging to its successors {**b,c,d**} (here all the heads and tails of **b**, **c** and **d**) ; the tail of **b** is preceded by **0** element of **{c,d}**; the tail of **c** is preceded by **1** element of **{d}** , and **0** is associated to the last palindrome **d** of **<abcd>**.

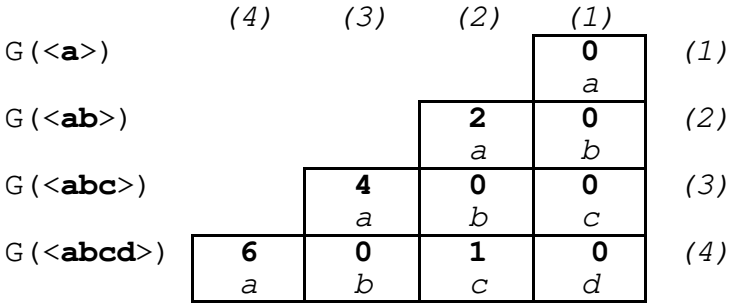
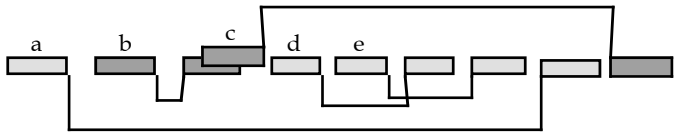
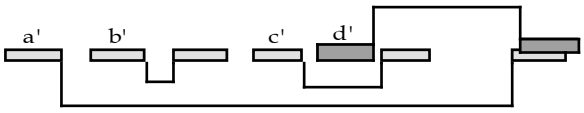



**Figure 5.** Dictionary encoding all contiguous substructures of the structure given in Figure 3. Indexes (in parenthesis) are mentioned on the right side of the lines and above the columns. The letter inside each cell is a comment that represents the name of the palindrome associated with the cell.




(a)

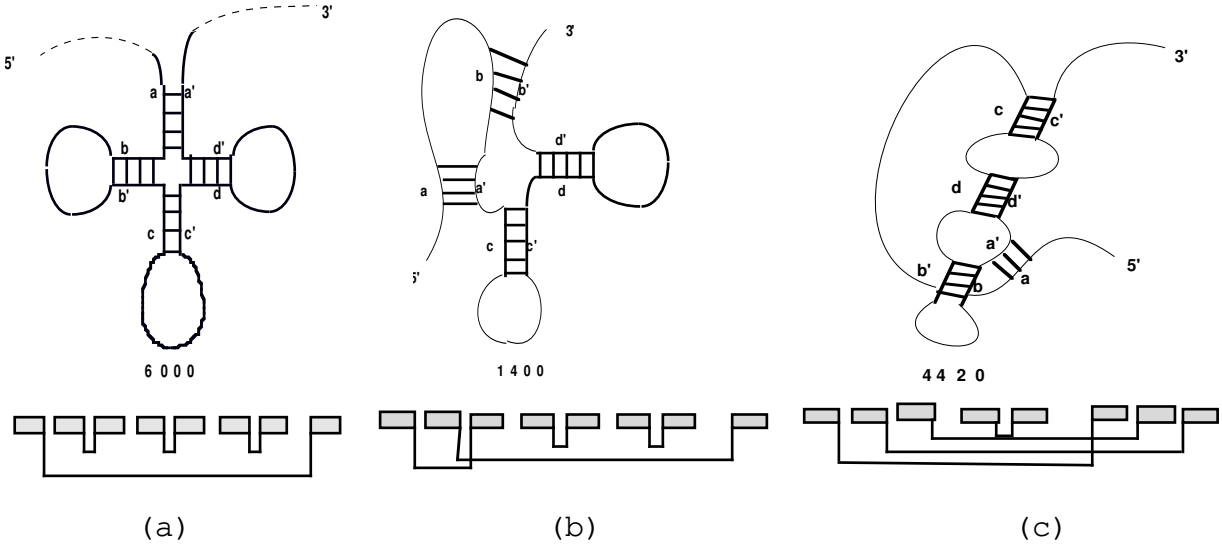



(b)

| | | | | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| | | | Du1 | | | | |
| 1 | | | 1 | | | | **0** *(a)* |
| 2 | | | 2 | | | **2** *(a)* | **0** *(b)* |
| 3 | | | 3 | | **4** *(a)* | **0** *(b)* | **0** *(d)* |
| 4 | | Ds1 | 4 | **6** *(a)* | **0** *(b)* | **1** *(d)* | **0** *(e)* |
| | | | Du2 | | | | |
| 5 | | | 1 | | | | **0** *(a)* |
| 6 | | | 2 | | | **1** *(a)* | **0** *(c)* |
| 7 | | | 3 | | **3** *(a)* | **2** *(c)* | **0** *(d)* |
| 8 | | | 4 | **5** *(a )* | **4** *(c)* | **1** *(d)* | **0** *(e)* |
| | D | | Du'1 | | | | |
| 9 | | | 1 | | | | **0** *(a')* |
| 10 | | | 2 | | | **2** *(a')* | **0** *(b')* |
| 11 | | Ds2 | 3 | | **4** *(a')* | **0** *(b')* | **0** *(c')* |
| | | | Du'2 | | | | |
| 12 | | | 1 | | | | **0** *(b')* |
| 13 | | | 2 | | | **0** *(b')* | **0** *(c')* |
| 14 | | | 3 | | **0** *(b')* | **1** *(c')* | **0** *(d')* |

(c)

**Figure 6.** An example of a general dictionary. Let **S** = {$\mathbf{s_1}$, $\mathbf{s_2}$} be a set of 2 sequences. $\mathbf{s_1}$ contains the set of palindromes {**a**, **b**, **c**, **d**, **e**} corresponding to the structural relations given in (a) (notice that **b** and **c** are incompatible, i.e. at least one of them does not correspond to a helix). We consider here the two virtual secondary structures of $\mathbf{s_1}$ : $\mathbf{u_1}$:<**abde**> and $\mathbf{u_2}$:<**abce**>.
$\mathbf{s_2}$ contains the set of palindromes {**a'**, **b'**, **c'**, **d'**} corresponding to the structural relations given in (b).The virtual secondary structures of $\mathbf{s_2}$ are $\mathbf{u'_1}$:<**a'b'c'**> and $\mathbf{u'_2}$:<**a'b'd'**>.
(c) shows the general dictionary D associated to **S** and encoding all structural patterns that occur either in $\mathbf{s_1}$ or $\mathbf{s_2}$. Given a quorum q = 2, the largest structural patterns occurring in all (2) sequences are represented by the 3-length repeated suffixes **4 0 0** and **0 1 0**.




(a) (b) (c)

**Figure 7**. Structural shapes of the three repeated structural patterns coded **6 0 0 0** (a), **1 4 0 0** (b) and **4 4 2 0** (c) (tRNA experiment).

---

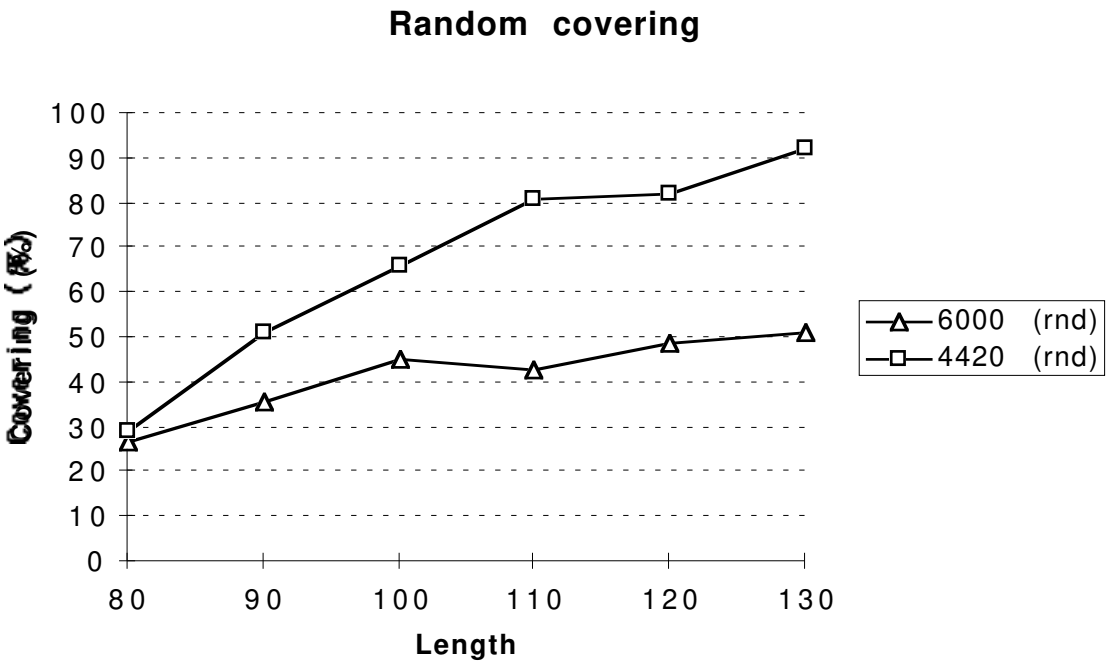
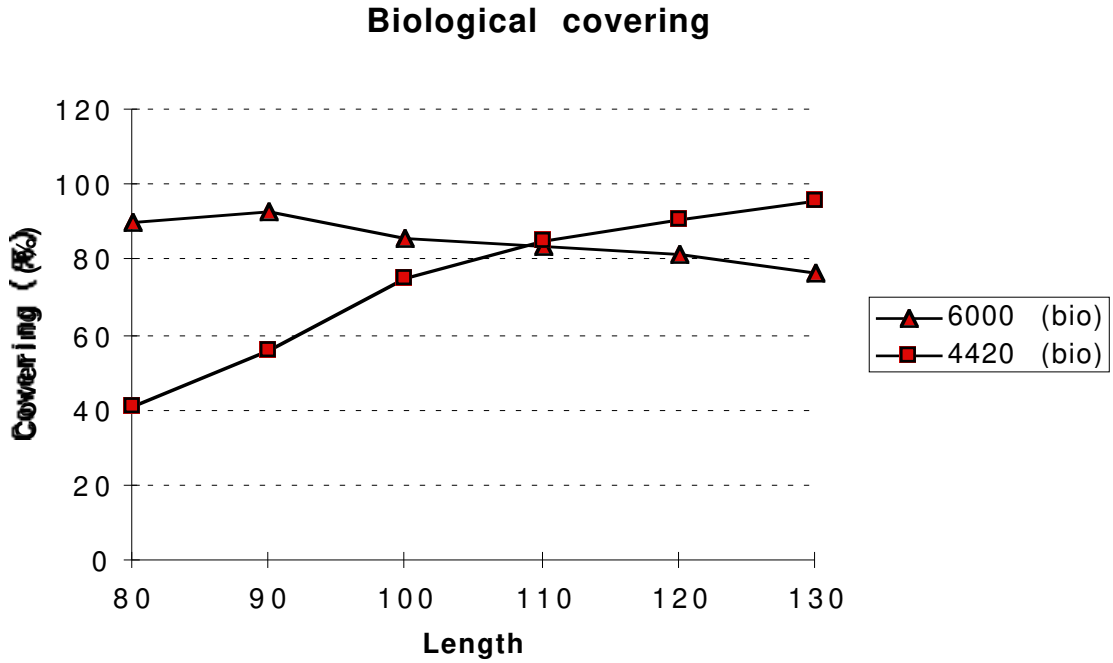



**Figure 8.** Random covering of the repeated structural patterns **6 0 0 0** and **4 4 2 0** (tRNA experiment) according to the length of the sequences.

---

**Figure 9.** Biological covering of the repeated structural patterns coded **6 0 0 0** and **4 4 2 0**




according to the length of the sequences (tRNA experiment). Note that here the biological covering weakly decreases. This is due to the following phenomena : as the first step we select a fixed number of structural patterns regardless of the length of the sequence, adding extra nucleotides sometimes result in missing relevant structural patterns.

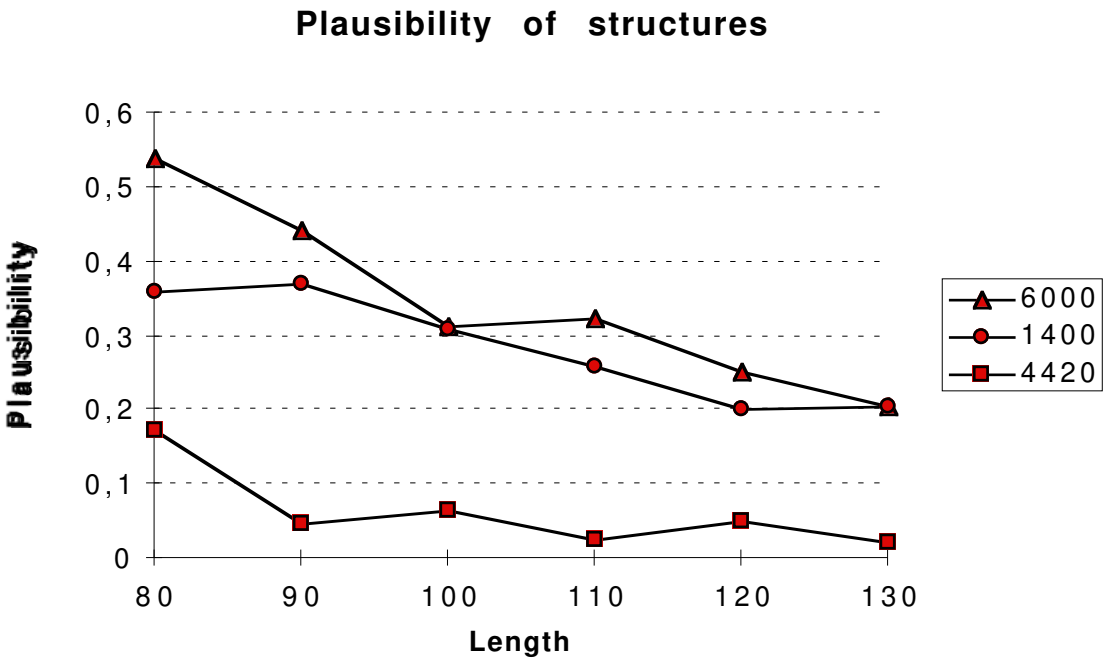



**Figure 10.** Plausibility of the repeated structural patterns **6 0 0 0**, **1 4 0 0** and **4 4 2 0** (tRNA experiment) according to the length of the sequences.