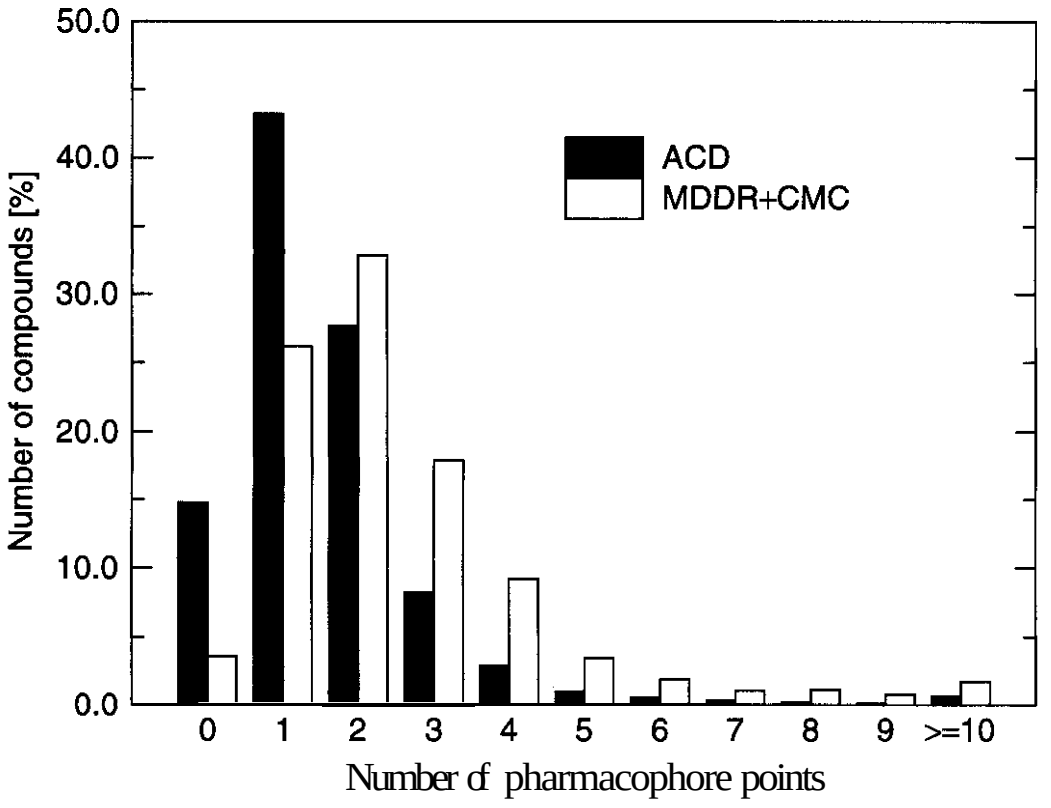

Figure 6.8. Number of **pharma**cophore points in drug databases (MDDR + CMC) and reagent **da**tabases (ACD).

- Pharmacophore points are fused and counted as one if they are separated by less than two carbon atoms.
- Molecules with less than two and more than seven pharmacophore points fail the filter.
- Amines are considered pharmacophore points but not azoles or diazines.
- Compounds with more than one carboxylic acid are dismissed.
- Compounds without a ring structure are dismissed.
- Intracyclic amines in the same ring are fused to one pharmacophore point.

The requirement of two distinct pharmacophore points neglects at least one very important class of drugs: biogenic amine-containing CNS drugs. Therefore, a second pharmacophore filter has been designed that requires only one pharmacophore point in small molecules of the type amine, amidine, guanidine, or carboxylic acid (PF2).

An analysis of drug databases and reagent-type databases reveal that about two thirds of drugs and nondrugs can be classified correctly by PF1. This performance is not as impressive as that of neural networks. However, as a filter for virtual screening, pharmacophore point filters offer some advantages. First, the occurrence and count of pharmacophore points can be evaluated on the building-block level of a virtual combinatorial library. No enumeration is necessary as for druglike neural nets. Second, the results of the pharmacophore point filter can be easily interpreted. Third, the settings of the filter can be easily adjusted (e.g., PF1 for non-CNS drugs, PF2 for CNS drugs).

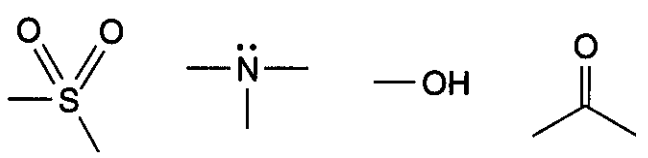

Figure 6.9. Functional motifs of drugs used to build pharmacophore points.

2.2 Focused Screening Libraries for Lead Identification

Without the knowledge about specific drug targets it is sometimes useful to apply virtual screening for the design of focused libraries of a few thousand compounds rather than to find a small number of hits to be tested against a specific target. To save resources it may sometimes be more prudent not to run the entire HTS file against a target protein; instead, a focused library with higher chances of containing hits may be scrutinized. Those focused libraries may be designed to target specific protein families such as GPCRs, kinases, or nuclear hormone receptors. They can also be enriched with privileged structures that occur