protein is calculated, providing a map of pseudo-affinities for each atom type or interaction type possibly present in the ligands to be docked. These maps then serve as look-up tables for the calculation of the interaction energy or scoring value during the docking process. Examples of docking programs using this approach are AutoDock (113–115), ICM (82, 116, 117), or ProDock (118, 119).

It should be noted that most of the mentioned representations of protein structure imply that the protein remains rigid during the docking process. As a matter of fact, docking under the assumption of a rigid protein is still common practice in standard applications. Although an acceptable simplification under certain circumstances, it can represent a serious limitation if only unbound protein structures are available. As a consequence, the inclusion of protein flexibility in the docking process is an active area of research, and a separate section is dedicated to this issue (cf. Section 3.2.1).

3.1.2 Ligand Handling. For the ligand, a complete representation in atomic coordinates is perfectly feasible. Ligand atoms may be used directly for matching with binding site descriptors or in the calculation of interaction energies in the case of energy-driven procedures. The central problem is conformational flexibility. Predicting the binding conformation of a ligand is in fact a major component of the docking problem, given that this conformation can significantly differ from that adopted in other environments.

Two general strategies for ligand handling may be distinguished: whole-molecule approaches and fragment-based methods. In the first case, the ligand is docked as an entire molecule. This is rather straightforward if the ligand is treated as a rigid body and only translational and rotational degrees of freedom are considered. Such rigid docking was common practice in early docking algorithms (106, 120). A straightforward extension to account for flexibility is to separately dock precalculated conformers of a given molecule (variant 1 in Fig. 7.2). Explicit docking of multiple conformers has, for example, been obtained with the FLOG program (121). FLOG deals with conformational flexibility by generating different conformers using distance geometry and docking each conformer in a rigid-body fashion. A similar approach has also been obtained with the DOCK program (122). To avoid redundancy in the docking, a common rigid fragment is identified, which is docked only once for the entire set of pregenerated conformers. The flexible portions of the molecule that determine the different conformations are subsequently scored based on the preplacement of the rigid fragment. Yet other examples for rigid docking of multiple conformers are provided by the programs FRED from OpenEye Scientific Software, which performs a fast exhaustive search over all possible orientations (123), and SYSDOC (124) or EUDOC (125), which use fast affine transformation to perform systematic searches over the translational and rotational degrees of freedom of the ligand.

Although this multi-conformer docking can be efficient and accurate for molecules with a limited number of discrete, low-energy conformations, it is less suited for larger and highly flexible molecules, simply because the number of possible conformations increases dramatically. Another way of partially accounting for conformational flexibility in whole-molecule rigid-body docking is to subject the initial matches to some kind of optimization that allows for conformational relaxation. This could be done with some standard energy minimization technique (126,127) or other procedures that resolve clashes of the initial placement by rotation about single bonds, as done, for example, in the docking program SLIDE (111).

A more rigorous treatment of ligand flexibility in whole-molecule docking is performed by sampling ligand conformation space during docking (variant 2 in Fig. 7.2). It normally requires ligand conformational energies to be evaluated besides intermolecular interaction energy. Molecular mechanics force fields are frequently applied for this purpose. Although a more exhaustive sampling of accessible conformations within the binding site is definitely achieved, an obvious disadvantage is the higher computational demand and possibly a reduced efficiency of the algorithm because of lengthy exploration of local minima.

An interesting variant of whole-molecule representations is the use of internal coordi-