mentioned that multistep hybrid docking procedures have been developed that combine rapid fragment-based searching with sophisticated MC or MD simulations (165, 166).

2.4.3 Scoring of Protein-Ligand Interactions. The problem of sampling the correct binding geometry (binding mode) of a protein-ligand complex is considered to be solved in many docking programs (167). However, to identify this correct binding mode by its lowest energy or score is a different matter; this is indeed the bottleneck of docking-scoring approaches today. The most important aspect of scoring functions for virtual screening is speed. Therefore, accuracy requirements are low; most functions used do not conceptually describe binding free energies. Therefore, these functions are typically not called energy functions but scoring functions. Three main scoring strategies are typically used in docking applications for virtual screening: force field scoring, empirical scoring, and knowledge-based scoring.

2.4.3.1 Force Field (FF) Scoring. Nonbonded interaction energy terms of standard force fields are typically used in FF scoring (e.g., in *vacuo* electrostatic terms; sometimes modified by scaling constants that assume the protein to be an electrostatic continuum) and van der Waals (vdW) terms (168–171). DOCK and GREEN (172) use the intermolecular terms of the AMBER energy function (173, 174), with the exception of an explicit hydrogen bonding term (147):

$$E_{\text{nonbond}} = \sum_{i}^{\text{lig}} \sum_{j}^{\text{prot}} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + 332 \frac{q_i q_j}{D r_{ij}} \right] \tag{6.1}$$

where each term is summed up over ligand atoms i and protein atoms j. A_{ij} and B_{ij} are the vdW repulsion and attraction parameters of the 6–12 potential, r_{ij} is the distance between atoms i and j, q is a point charge at each of the atoms, and D is the dielectric constant. Intraligand interactions are added to the score. Up to a 100-fold gain in docking time can be achieved by precomputing these terms on a 3D grid that represents the protein during docking (155, 175). More recently, solvation terms have been added to FF scores. Examples include generalized Born/surface area approaches (176) or atomic solvation parameters (177–179).

2.4.3.2 Empirical Scoring. Empirical scoring functions are multivariate regression methods. They fit coefficients of physically motivated contributions to binding free energy in reproduction of measured binding affinities of a training set of protein-ligand complexes with known 3D structure. As an example, the docking program FlexX (180) uses a scoring function similar to that of Bohm (181, 182). It calculates the sum of free-energy contributions from the number of rotatable bonds in the ligand, hydrogen bonds, ion-pair interactions, hydrophobic and pi-stacking interactions of aromatic groups, and lipophilic interactions:

$$\begin{aligned} \Delta G = {} & \Delta G_0 + \Delta G_{\text{rot}} N_{\text{rot}} \\ & + \Delta G_{\text{hb}} \sum_{\text{neutral\_Hbonds}} f(\Delta R, \Delta\alpha) \\ & + \Delta G_{\text{io}} \sum_{\text{ionic\_int}} f(\Delta R, \Delta\alpha) \\ & + \Delta G_{\text{aro}} \sum_{\text{aro\_int}} f(\Delta R, \Delta\alpha) \\ & + \Delta G_{\text{lipo}} \sum_{\text{lipo.cont}} f^*(\Delta R) \end{aligned} \tag{6.2}$$

where ΔG_0, ΔG_{rot}, AG,, ΔG_{io}, ΔG_{aro}, and ΔG_{lipo} are adjustable parameters that are fitted; $f(\Delta R, \Delta\alpha)$ is a scaling function penalizing deviations from the ideal geometry; and N_{rot} is the number of freely rotatable bonds. The interaction of aromatic groups is an addition to Bohm's original force-field design (181, 182). The lipophilic contributions are calculated as a sum of atom-pair contacts in contrast to evaluating a surface grid as in Bohm's scoring function. Bohm's scoring function and its FlexX implementation are being improved and additional terms are being tested (see, e.g., Refs. 182 and 183).

2.4.3.3 Knowledge-Based Scoring. Because the forces that govern protein-ligand interactions are so complex, an implicit approach to capture all relevant terms of protein-ligand