

Synergistic Use of Virtual Screening and Biophysical Methods for the Protein-Based Design of Peptidomimetics

Didier Rognan^{a,b*}, Caterina Bissantz^{a,b}, Séverine Dédier^a, Antoine Logean^{a,b}, and Stefan Reinelt^a

Abstract: Combining virtual screening with biophysical studies of protein-ligand complexes is an effective tool for designing new peptidomimetics. When a three-dimensional structure of the target protein is known, automated docking of chemical databases can be used as a powerful filter to reduce the number of molecules that will be further tested. Easy screening of potential hits can then be performed using fluorescence polarization techniques, if a fluorescent-labeled ligand already exists for the target. In addition, thermodynamic properties of protein-ligand complexes can be measured by circular dichroism spectroscopy.

Keywords: Circular dichroism · Fluorescence polarization · Peptidomimetic · Pharmaceutical chemistry · Virtual screening

Pharmaceutical chemistry has just entered a new era with the complete sequencing of a human genome [1]. The number of potentially interesting drug targets will no doubt dramatically increase in the coming years. To just give an idea of the revolution to come, about 500 targets are currently used by the modern pharmacopoeia, but between 5000 and 10 000 are foreseen in the near future [2]. A significant number of these new targets will probably be orphan targets for which no or very little information on endogenous ligands will be available. To get first insights into the molecular structures of such ligands, peptide combinatorial libraries [3] offer the great advantage of being synthetically accessible within short time scale so that peptide

ligands often constitute the very first step in the design of novel pharmaceuticals [4]. As peptides remain very sensitive to enzymatic degradation, they are, however, not well suited for clinical development. One intense research area of the last decade has therefore been the design of peptidomimetics sharing binding properties similar to that of their peptide parents but with enhanced pharmacokinetic profiles. Significant success had been reported notably in the field of protease inhibitors [5][6] but these are often restricted to short-length peptide mimics (about four to six residues). Mimicking longer peptide sequences is presently possible in only a very few cases [7].

1. Virtual Screening as a Potent Hit-finding Method

1.1. How to Begin with Virtual Screening

We have been interested for the last two years in developing a rational design strategy aimed at mimicking long peptide sequences (about ten amino acids) by small molecular weight organic molecules. Ultra-fast high-throughput screening (uHTS) of huge collections of chemicals is becoming standard in major pharmaceutical companies [8] but is far beyond the means of any academic envi-

ronment. Furthermore, its real relevance to the design of better clinical candidates is still a matter of debate [9][10]. Progress in computer science has recently opened new avenues for integrating cheminformatics, data mining and computational chemistry in lead finding. Several large two-dimensional databases (corporate or academic) comprising several hundred thousand molecules are now available and can be filtered prior to experimental screening. Starting from the Advanced Chemicals Directory (MDL, San Leandro) which currently contains about 290 000 chemicals, a 'chemical filtering' enables us to discard about 75% of the database using very simple filters: (i) molecular weight under 500, (ii) eliminating chemically reactive and toxic molecules, (iii) discarding inorganic compounds (Fig. 1).

A 'pharmacokinetic' filter might then be applied to retrieve 'drug-like' molecules sharing specific physicochemical properties (number of H-bond donors and acceptors, logP, polar surface area) [11][12] that maximize oral bioavailability. For each of the 75 000 remaining molecules, three-dimensional (3D) coordinates can be accurately computed from the two-dimensional (2D) structure using 3D conversion programs like CORINA [13]. As a last filter, the 3D database is docked into the binding site of the target

*Correspondence: PD. Dr. D. Rognan^a

^aDepartment of Applied Biosciences
ETH Zürich
Winterthurerstrasse 190
CH-8057 Zürich

Tel.: +41 1 635 60 36

Fax: +41 1 635 68 84

E-Mail: didier@pharma.ethz.ch

^bNew address from Oct 1st, 2000:

Laboratoire de Pharmacochimie de la
Communication Cellulaire
UMR CNRS 7081

74, route du Rhin – B.P. 24

F-67401 Illkirch, France

Tel.: +33 3 88 67 68 07

Fax: +33 3 88 67 47 94

E-Mail: didier.rognan@pharma.u-strasbg.fr

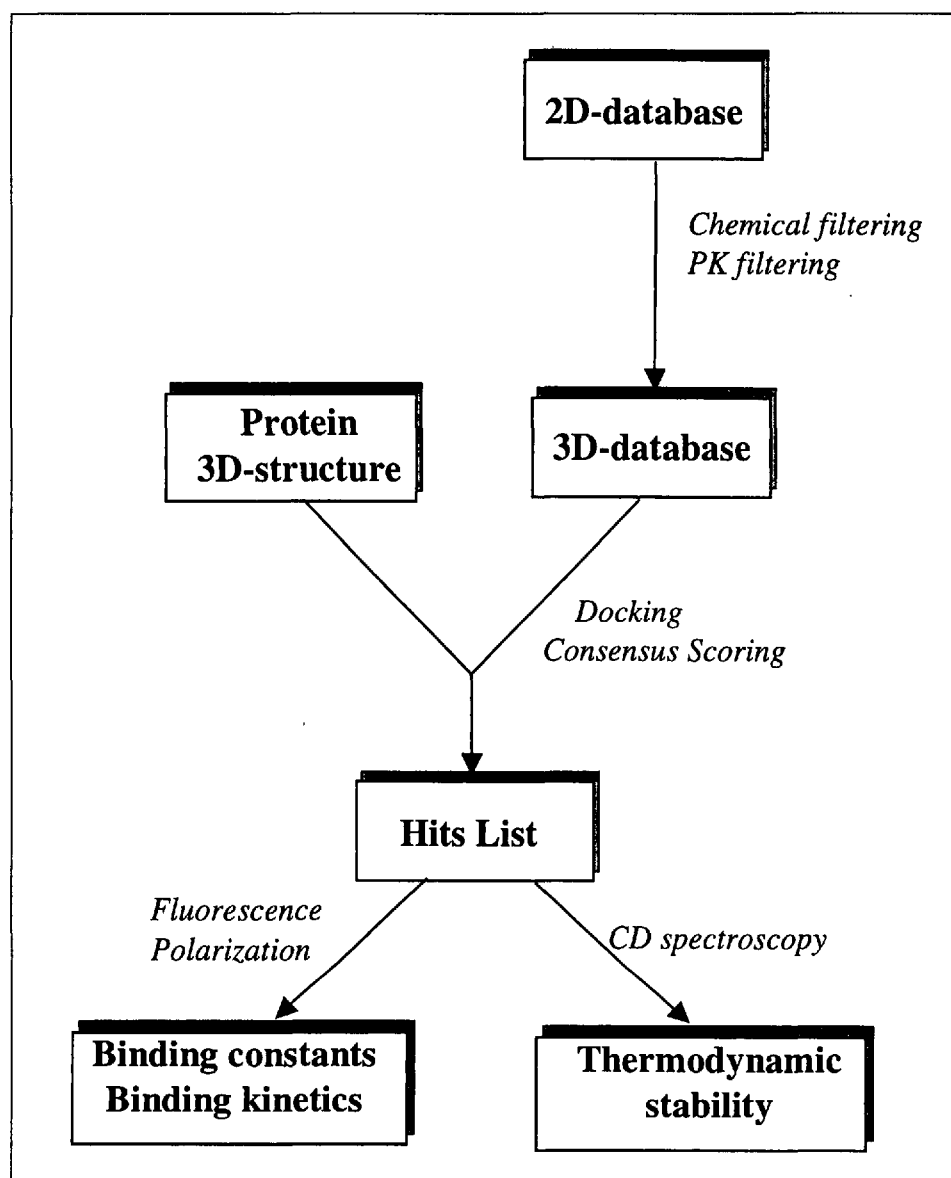


Fig. 1. Our approach for integrating protein-based virtual screening with biophysics of protein-ligand interactions: This procedure is currently used for two main projects: (i) design of non-peptide T-cell receptor antagonists, (ii) design of phosphoenolpyruvate:sugar phosphotransferase Enzyme I inhibitors.

protein to retrieve molecules presenting optimal complementarity (score) to the target's active site. This implies that (i) the 3D structure of the target protein has been solved by either X-ray diffraction and/or NMR spectroscopy, (ii) in absence of any experimentally-determined protein structure, a high-resolution protein model should be available, (iii) the binding site has been experimentally determined.

1.2. Virtual Screening: Reality or Illusion?

Our experience in database docking strongly suggests the simultaneous use of several docking and scoring techniques (Fig. 2). We critically judged the accuracy of several docking/scoring schemes by trying to find ten true hits out of a random

library of 1000 'drug-like' molecules for two protein test cases. The thymidine kinase (TK) of the Herpes Simplex Virus was first chosen as a hard test because of its open active site, a significant degree of induced fit upon ligand binding, the participation of water molecules in mediating ligand binding. The antagonist-bound conformation of the estrogen receptor (ER) was chosen as an easier second test. Its active site is made of a close apolar cavity, few conformational changes of the protein upon ligand binding are observed. In both cases, sufficient experimental data (binding free energies of protein-ligand complexes, X-ray structures) were available to evaluate the accuracy of virtual screening methods. It turned out that even in the very difficult TK case, it was possible to (i) discrimi-

nate true hits from random ligands (Fig. 2), (ii) enrich a reduced library by at least a factor 10 (Fig. 3), (iii) retrieve about 50% of all true hits.

However, significant progress still needs to be made in order to (i) predict the exact protein-bound orientation of the ligands, (ii) predict the absolute binding free energy of the retrieved hit, (iii) discriminate true hits from chemically similar but inactive compounds, (iv) catch 100% of the true hits. Interestingly, promising preliminary results have been obtained by screening chemical databases against a homology model of a G Protein-Coupled receptor, suggesting that virtual screening does not absolutely require high-resolution X-ray protein structures.

Thus, virtual screening programs can be used with significant efficiency in hit hunting programs for prioritizing ligand synthesis and experimental testing. We are currently applying virtual screening methods to develop totally new immunosuppressants as well as to discover non-peptide inhibitors of bacterial enzymes (Enzyme I of the phosphoenolpyruvate:sugar phosphotransferase) for a new class of antibiotics. In the latter case, virtual screening of a large chemical database (270 000 compounds) allowed a list of 20 potential hits to be proposed within four days, one of which displayed μM affinity for the target enzyme. Scoring hits as accurately as possible being an essential feature of virtual screening, we developed in collaboration with Novartis a new scoring function [14] particularly suited to predict absolute binding free energies from 3D models.

2. Biophysical Studies of Protein-Ligand Interactions

Once a hit list has been proposed from virtual screening, potentially interesting molecules should be easily tested for binding affinity to their host target. We currently use two main techniques (fluorescence polarization, circular dichroism) for this purpose.

2.1. Fluorescence Polarization: a Simple and Fast Method for Studying Protein-Ligand Complexes in Solution

A main disadvantage of most methods aimed at measuring binding constants is that they require separation of the bound from the free ligand. Using fluorescence polarization (FP), no washing and separation steps are required. By

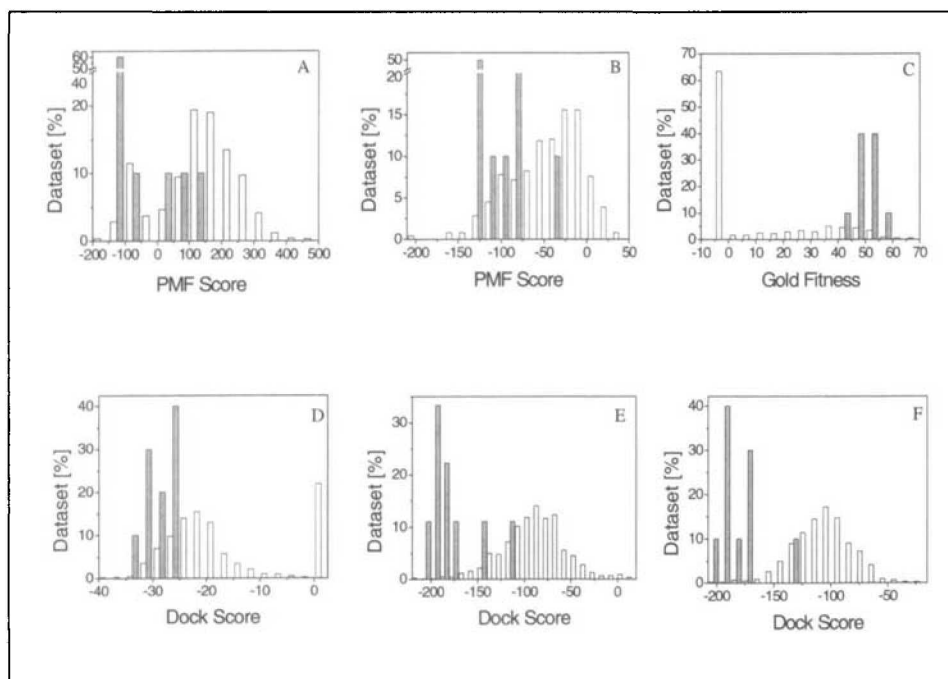


Fig. 2. Virtual screening of a 3D database containing 990 random ligands and ten true hits. Ligands were flexibly docked using either Dock4.0 (left panels) or FlexX1.8 (middle panels) and Gold1.1 (right panels) at a pace of about 100 sec/molecule. All potential hits were ranked using seven different scoring functions (Chemscore, Dock, FlexX, Fresno, Gold, Pmf and Score) against two different targets (Herpes Simplex virus thymidine kinase or TK, estrogen receptor or ER). The best docking/scoring combinations are shown for TK (A-C) and ER screening (D-F). Scores of true hits and 'random' molecules are indicated by grey and white bars, respectively.

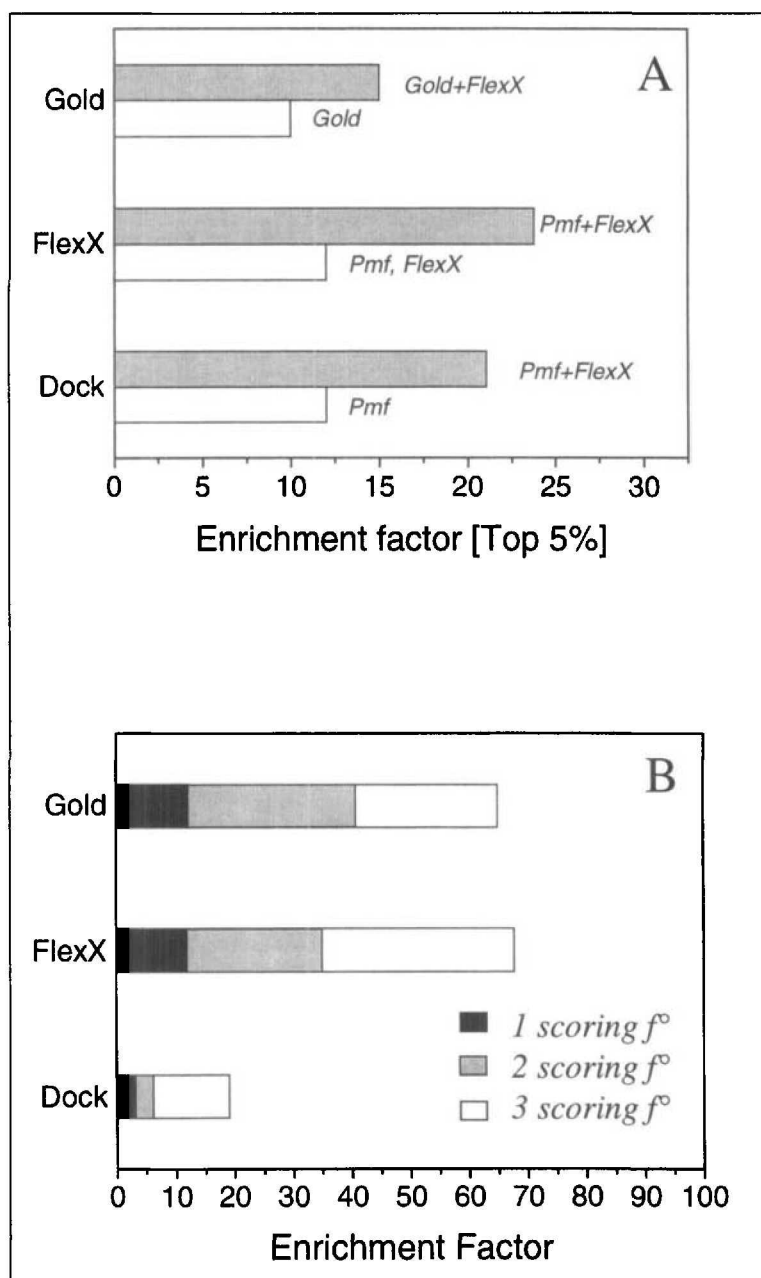


Fig. 3. Effect of consensus scoring on the enrichment factor in true hits obtained by virtual screening of TK and ER. The enrichment factor is the ratio of the percentage of true hits found by virtual screening to the same percentage found by random screening. The best enrichment factors are shown for TK screening (panel A) whereas averaged enrichment factors are displayed for ER screening (panel B) using either a single scoring function (dark bars) or a consensus scoring with two (grey bars) or three (white bars) scoring functions.

simply mixing a soluble preparation of the target protein, a fluorescent-labeled ligand with or without a competitor, one can easily determine binding constants.

Fluorescence polarization (FP) was originally described in 1926 by Perrin [15] (Eqn. 1)

$$FP = (I_{\parallel} - I_{\perp}) / (I_{\parallel} + I_{\perp}) \quad (1)$$

I_{\parallel} : Fluorescence intensity in the parallel plane

I_{\perp} : Fluorescence intensity in the perpendicular plane.

A fluorescent-labeled molecule, when excited with plane polarized light, will emit light in the same plane if the molecule slowly rotates throughout the excited state (a few ns in the case of fluorescein). If the molecule tumbles much faster, light will be emitted in all directions. Thus, the degree to which emission intensity moves from the original plane to the perpendicular one is related to the mobility of the probe. If the probe is free in solution, FP values will be rather low (typically around 50 mP). If the probe is bound to a macromolecule, fluorescence emission will be polarized and FP values will be much higher (up to 500 mP).

We routinely use FP for studying class I MHC-ligand interactions (Fig. 4) to determine (i) association and dissociation kinetic constants, (ii) half-lives ($t^{1/2}$) of MHC-ligand complexes, (iii) screening potential new hits, (iv) determine binding constants (K_i , IC_{50}).

2.2. Folding/Unfolding of Protein-Ligand Complexes Monitored by CD Spectroscopy

Determining kinetic parameters is essential in the field of MHC ligands as the stability of the MHC-ligand complex is much more important than the simple binding affinity of the ligand for its host MHC protein. Thus determining the thermodynamic stability of protein-ligand complexes is a very important issue. One method of choice is to follow by CD spectroscopy the loss of the protein's secondary structure upon denaturation using either denaturing agents (urea, guanidinium chloride) or temperature. We typically use the temperature denaturation method, as it is less time-consuming than other approaches. The temperature at which 50% of the protein unfolds (indicated as melting temperature or T_m) will be dependent on the thermodynamic stability of the protein-ligand complex. A typical denaturation curve of a class I MHC-peptide complex is indicated in

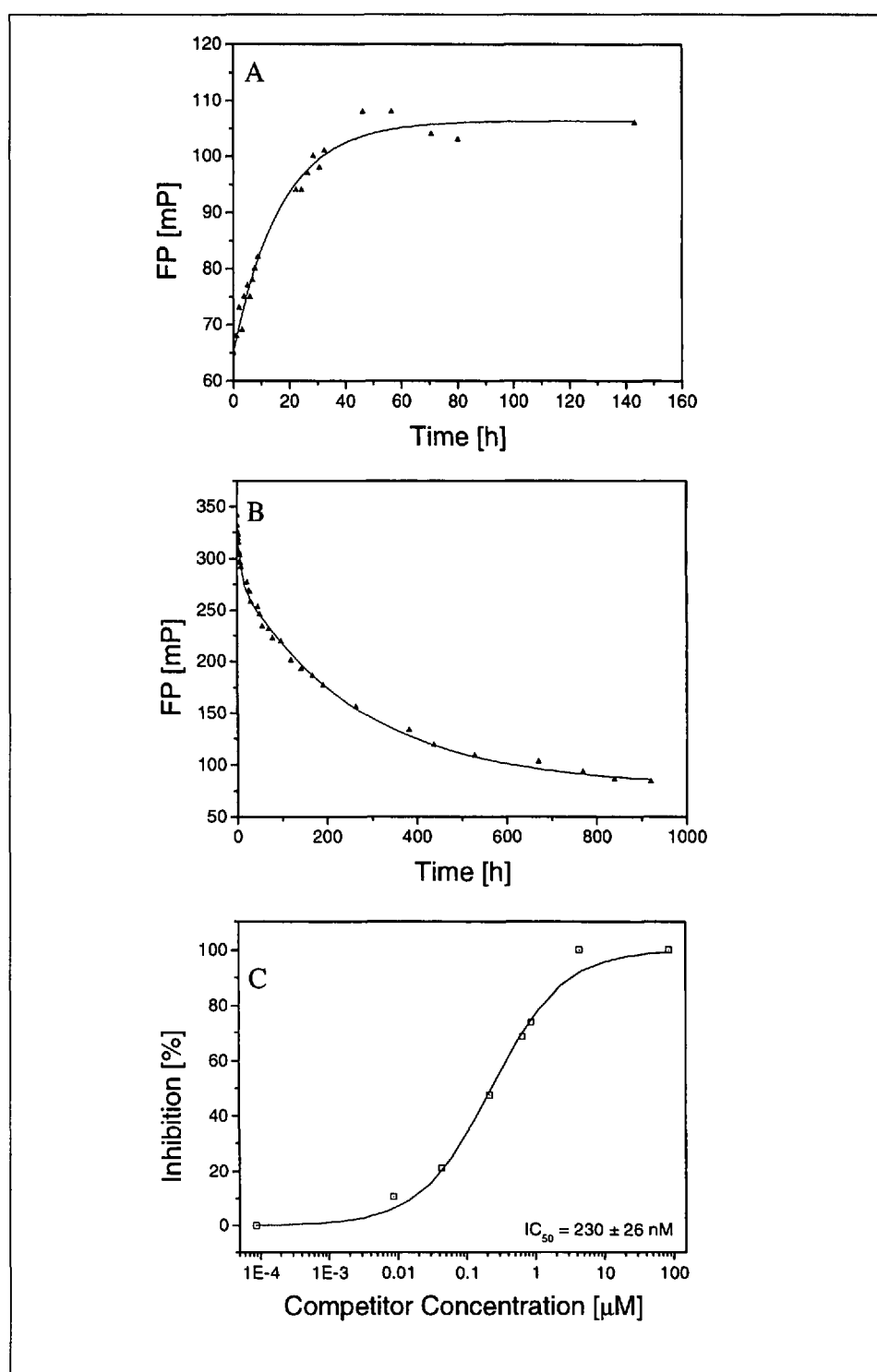


Fig. 4. Fluorescence polarization studies of peptide association/dissociation to a class I MHC protein. A. Kinetics of peptide association, (B) kinetics of peptide dissociation, (C) determination of binding constants (IC_{50} , K_i) for a peptide competitor.

Fig. 5a. As the temperature increases, the peptide ligand is progressively expelled from the binding site and the free protein unfolds, thus leading to a loss in secondary structure. This method has the advantage that it requires no labeling of either the protein or the ligand. Melting temperatures (T_m) of protein-ligand complexes are usually strongly correlated to the equilibrium dissociation constant of the

ligand (K_D) [16] and in some cases to experimental inhibition constants (K_i). Thus small changes of either the ligand or the protein may be easily monitored [17] from the thermodynamic point of view (Fig. 5B). One particular interesting aspect of such experimental comparisons is that ΔT_m values measured by CD spectroscopy can be related to binding free energy differences [18].

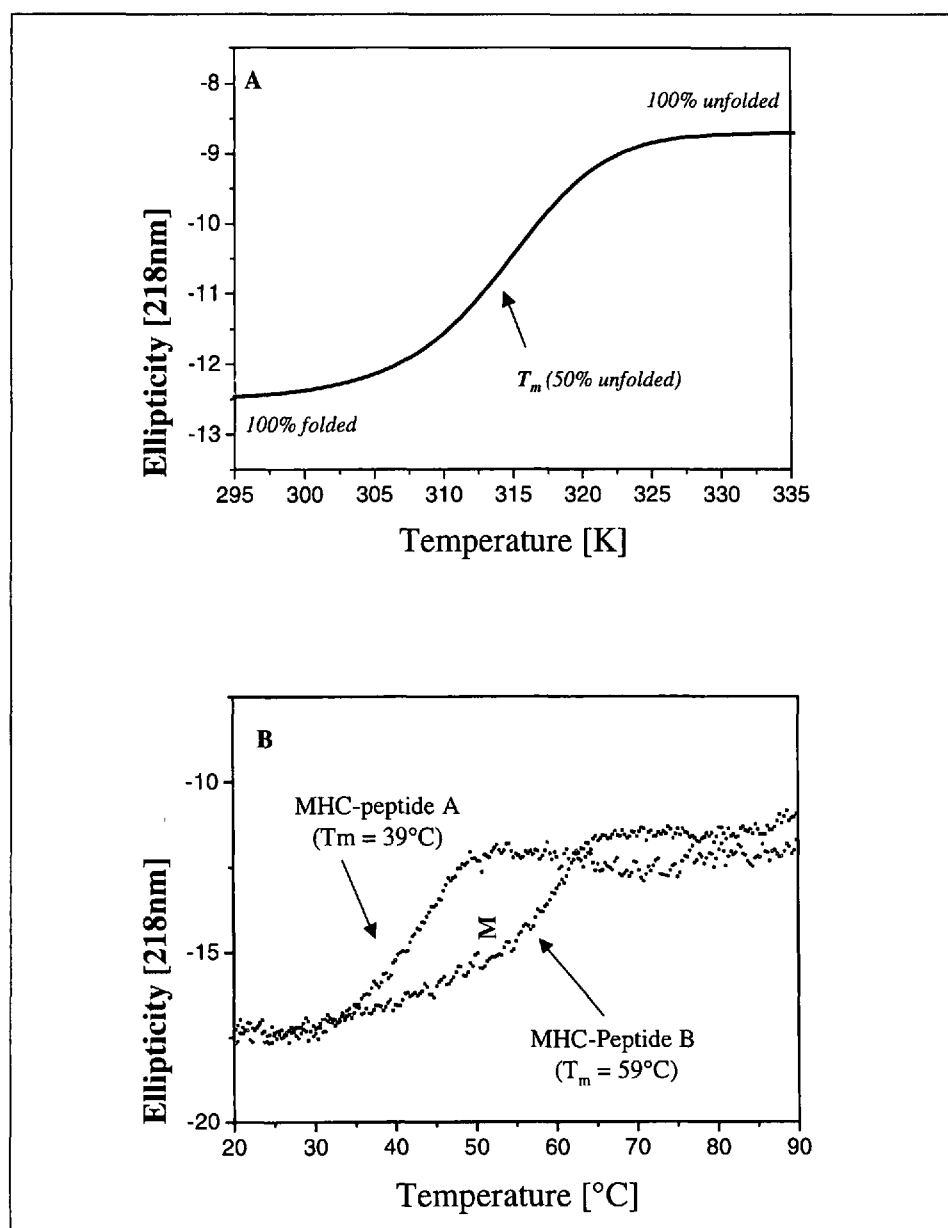


Fig. 5. Thermal denaturation of MHC-peptide complexes monitored by CD spectroscopy at 218 nm [17]. (A) Typical denaturation curve, (B) Comparison of two complexes between a class I MHC protein and two peptides (peptide A: $T_m = 39^\circ\text{C}$; peptide B: $T_m = 59^\circ\text{C}$)

3. Outlook

In the very near future, the number of therapeutically interesting targets will increase tremendously. More than ever, there will be a need to rationalize experimental screening of potential hits to decrease as much as possible the size of chemical libraries to be screened. Virtual screening methods are now in a position to be used as a filter prior to experimental testing. They have already been successful, notably on orphan targets for which very little information about the endogenous ligand is known. As pharmacoki-

netic parameters also begin to be taken into account for filtering libraries of 'drug-like' molecules, coupling virtual to experimental screening shows a very promising approach for hunting not only hits but also lead compounds with better pharmacokinetic profiles.

Acknowledgments

The authors wish to thank the Swiss National Science Foundation (Project No. 31-57307-99) and ARPIDA AG (Münchenstein, Switzerland) for financial support.

- [1] <http://www.pcorporation.com/press/prccorp040600.html>.
- [2] J. Drews, *Science* **2000**, 287, 1960.
- [3] G. Jung, A. Beck-Sickinger, *Angew. Chem. Int. Ed. Engl.* **1992**, 31, 367.
- [4] J. Huff, *J. Med. Chem.* **1991**, 8, 2305.
- [5] P.Y. Lam, P.K. Jadhav, C.J. Eyermann, C.N. Hodge, Y. Ru, L.T. Bachelier, J.L. Meek, M.J. Otto, M.M. Rayner, Y.N. Wong, *Science* **1994**, 263, 380.
- [6] P.E. Sanderson, T.A. Lyle, K.J. Cutrona, D.L. Dyer, D.B. Dorsey, C.M. McDonough, A.M. Naylor-Olsen, I.W. Chen, Z. Chen, J.J. Cook, C.M. Cooper, S.J. Gardell, T.R. Hare, J.A. Krueger, S.D. Lewis, J.H. Lin, B.J. Lucas Jr., E.A. Lyle, J.J. Lynch Jr., M.T. Stranieri, K. Vastag, Y. Yan, J.A. Shafer, J.P. Vacca, *J. Med. Chem.* **1998**, 41, 4466.
- [7] B. Zhang, G. Salituro, D. Szalkowski, Z. Li, Y. Zhang, I. Royo, D. Vilella, M.T. Diez, F. Pelaez, C. Ruby, R.L. Kendall, X. Mao, P. Griffin, J. Calaycay, J.R. Zierath, J.V. Heck, R.G. Smith, D.E. Moller, *Science* **1999**, 284, 974.
- [8] T. Mander, *Drug Discov. Today* **2000**, 5, 223.
- [9] R. Lahana, *Drug Discov. Today* **1999**, 4, 447.
- [10] C.S. Ramesha, *Drug Discov. Today* **2000**, 5, 43.
- [11] S.D. Pickett, I.M. McLay, D.E. Clark, *J. Chem. Inf. Comput. Sci.* **2000**, 2, 263.
- [12] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feney, *Drug Delivery Rev.* **1997**, 23, 3.
- [13] J. Gasteiger, C. Rudolph, J. Sadowski, *Tetrahedron Comp. Method.* **1990**, 3, 537.
- [14] D. Rognan, S.L. Laumoeller, A. Holm, S. Buus, V. Tschinke, *J. Med. Chem.* **1999**, 42, 4650.
- [15] F. Perrin, *J. Phys. Radium* **1926**, 1, 390.
- [16] C.S. Morgan, J.M. Holton, B.D. Olafson, P.J. Bjorkman, S.L. Mayo, *Protein Sci.* **1997**, 6, 1771.
- [17] S. Dédier, S. Reinelt, T. Reitingner, G. Folkers, D. Rognan, *J. Biol. Chem.* **2000**, 275, 27055.
- [18] M. Bouvier, D.C. Wiley, D. C. *Science* **1994**, 265, 398.