

Objektyp: **BackMatter**

Zeitschrift: **Horizonte : Schweizer Forschungsmagazin**

Band (Jahr): - **(2006)**

Heft 70

PDF erstellt am: **26.09.2024**

Nutzungsbedingungen

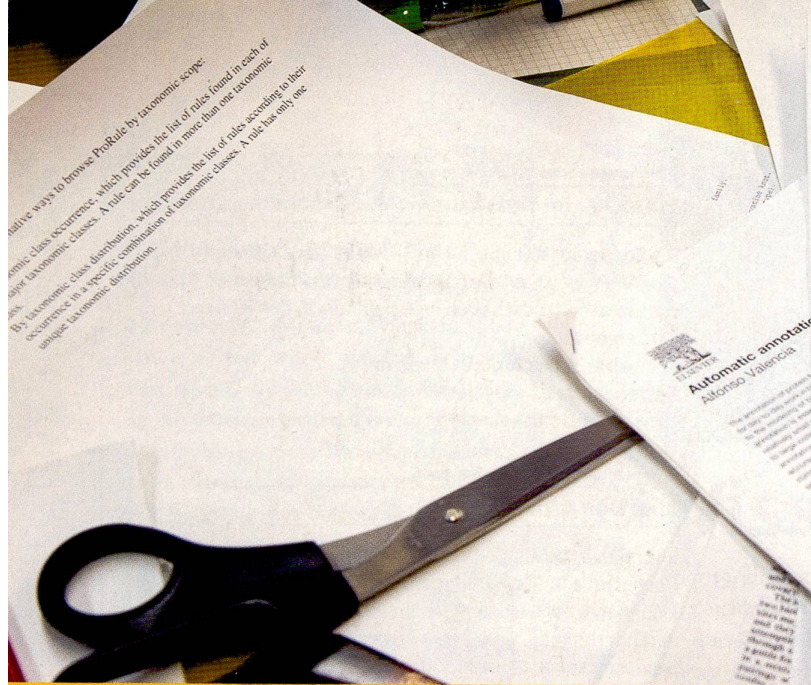
Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

<http://www.e-periodica.ch>



A Method for Localizing Ligand Binding Pockets in Protein Structures

Fabian Glaser,¹ Richard J. Morris,¹ Rafael J. Nijmawati,¹ Roman A. Laskowski,¹ and Janet M. Thornton^{1,2}
¹European Bioinformatics Institute, European Molecular Biology Laboratory, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, United Kingdom
²Department of Biochemistry and Molecular Biology, University College, London, United Kingdom

ABSTRACT The accurate identification of ligand binding sites in protein structures can be valuable in determining protein function. Once the binding site is known, it becomes easier to perform *in silico* and experimental procedures that may allow the ligand type and the protein function to be determined. For example, binding pocket shape analysis relies heavily on the correct identification of the ligand binding site. We have developed SURFNET, a novel three-stage method for identifying the location and shape of potential ligand binding pockets in protein structures. In the first stage, the SURFNET program identifies clefts in the protein surface that are potential binding sites. In the second stage, these clefts are trimmed in size by cutting away regions distant from highly conserved residues, as defined by the ConSurf-IHSSP database. The largest clefts that remain tend to be those where ligands bind. To test the approach, we analyzed a nonredundant set of 244 protein structures from the PDB and found that SURFNET ConSurf identifies a ligand binding pocket in 25% of them. The trimming procedure reduces the original cleft volume by 30% on average, while still encompassing an average 87% of the ligand volume. From the analysis of the results we conclude that for those cases in which the ligands are found in large, highly conserved clefts, the combined SURFNET ConSurf method gives pockets that are a better match to the ligand shape and location. We also show that this approach works better for enzymes than for nonenzyme proteins. *Proteins* 2000,42:479–488. © 2000 Wiley-Liss, Inc.

Key words: SURFNET; ConSurf-IHSSP; residue conservation; ligand identification; ligand shape; protein clefts; functional sites

INTRODUCTION

The prediction of protein function from 3D structure is becoming increasingly important, as more and more genes structures of unknown function are added to the various structural genomics projects (SGP).^{1–3} In this context there have been many attempts to use structural information, particularly features relating to the shape and volume of potential binding sites, as a means of determining function. Some of the shape, size, and structural composition

of the protein surface dictate the type of interaction the protein can make with its cognate ligand or other interacting partner (DNA, a second protein, etc.). In many studies, a variety of computational methods^{4–11} have been used to predict the location and shape of potential binding sites. These methods have concentrated on the analysis of those proteins where the ligand type is known. The overall objective is to predict, as accurately as possible, the ligand location (binding site), ligand type (cognate ligand, cofactor, etc.), and ultimately the protein function, based on the structure of that protein. This presents a number of important difficulties. For example, in most cases the binding site pocket is considerably larger than the ligand itself. Furthermore, in many enzymes the binding pocket is occupied by more than one molecule (e.g., substrate and cofactor).

Previous approaches have aimed to shed light on these issues, mainly by identifying the approximate binding-site region by describing and comparing binding-site properties in an indirect way of comparing functions. For example Mann et al.¹² dissect protein pockets into small binding volumes coupled with their physico-chemical characteristics (i.e., hydrophobicity, charge, etc.) and then try to map them to potential binding partners. Silverstein et al.¹³ developed a method to identify substrate binding sites in proteins using computational solvent mapping. This method identifies pockets on the protein surface in which organic molecules tend to cluster, taking into account the free energy of binding of those probes to the protein residues. Regions in which several different types of solvent probes bind, are predicted to be ligand-binding sites. Schmidt et al.¹⁴ used a different approach to detect similarities between a query binding site and other similar preprocessed protein cavities stored in the Cavbase data-

Grant sponsor: the Biotechnology Research Council (BRC); Grant sponsor: Structural Proteomics in Europe (SPPE); Grant number: 4815-CT-2002-00090; EMBL contract number; Grant sponsor: the Wellcome Trust; Grant number: 058113 (J.M.T.).

Rafael J. Nijmawati's present address is John Innes Centre, Norwich, Norfolk, UK. Contact: R.J.Nijmawati@jic.ac.uk

Correspondence to: Fabian Glaser, EMBL EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB3 0EH, United Kingdom. E-mail: fglaser@ebi.ac.uk

Received 21 April 2000; Received 14 July 2000; Accepted 15 August 2000

Published online 21 December 2000 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.10000

