

Zeitschrift: Horizonte : Schweizer Forschungsmagazin
Herausgeber: Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen
Forschung
Band: - (2006)
Heft: 70

Endseiten

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

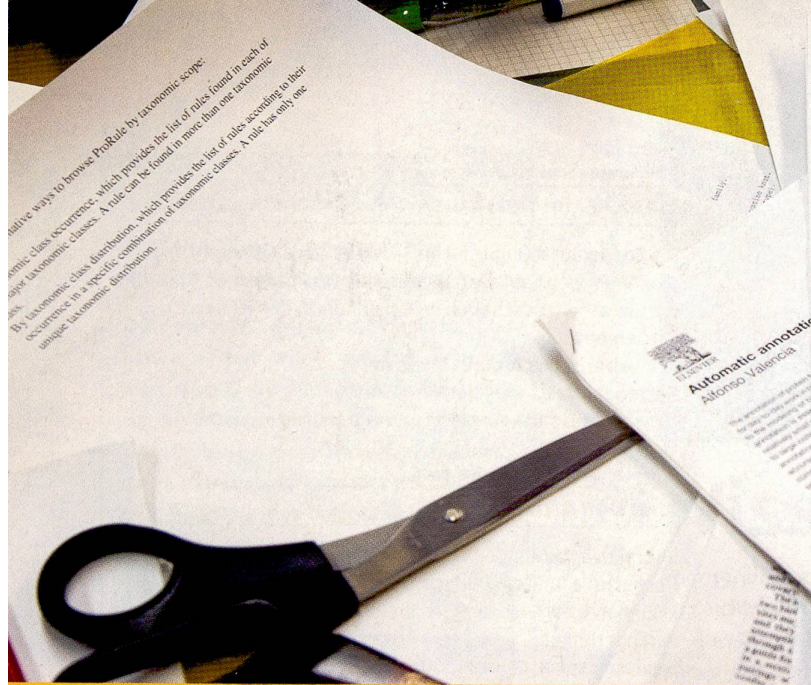
L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 12.05.2025

ETH-Bibliothek Zürich, E-Periodica, <https://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 Hiron A. Lashkovsk,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 can be used to predict the localization of the ligand binding site. We have developed SURFNET, a modular software 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-HSSP database, 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 75% of them. The trimming procedure reduces the original cleft volumes 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-HSSP; 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 protein structures of unknown function are added to the various structural genomics projects (SGP).^{1–5} In this context there have been many attempts to use structural information, particularly features relating to the functional sites of proteins,^{6–10} as a means of determining function. Since the shape, size, and chemical 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.), many studies have concentrated on the analysis of these properties using a variety of computational methods.^{11–14} 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, mostly by identifying the approximate binding-site region or by describing and comparing binding-site properties (as an indirect way of comparing functions). For example, Mason et al.¹⁵ 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.¹⁶ 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 grooves to the protein residues. Regions, in which several different types of solvent probes bind, are predicted to be ligand-binding sites. Schmidt et al.¹⁷ used a different approach to detect similarities between a query binding site and other similar preprocessed protein cavities stored in the Carbone data-

Grant sponsor: the Structural Genomics Consortium (SGC); Grant sponsor: European Molecular Biology Laboratory (EMBL); Grant sponsor: Structural Genomics in Europe (SGPE); Grant number: 1815-CT-2000-00050; ECU contract number; Grant sponsor: the Biotechnology and Biological Sciences Research Council (BBSRC); Grant number: 13863-CT-2000-00050.

Richard J. Morris's present address is John Innes Centre, Norwich Bioscience Park, Colney, NR3 2TH, Norwich, United Kingdom.
Correspondence to: Fabian Glaser, EMBL EBI, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SD, United Kingdom. E-mail: fglaser@ebi.ac.uk

Received 21 April 2000; Received 14 July 2000; Accepted 15 August 2000
Published online 21 November 2000 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20000

© 2000 WILEY-LISS, INC.