Protein kinase resource: An integrated environment for phosphorylation research

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · APRIL 2006

Impact Factor: 2.63 · DOI: 10.1002/prot.20825 · Source: PubMed

6 AUTHORS, INCLUDING:

Nina M Haste
University of California, San Diego
22 PUBLICATIONS **1,377** CITATIONS
SEE PROFILE

Michael Gribskov
Purdue University
101 PUBLICATIONS **10,744** CITATIONS
SEE PROFILE

Protein Kinase Resource: An Integrated Environment for Phosphorylation Research

Roland H. Niedner,[1,*] Oleksandr V. Buzko,[1,*] Nina M. Haste,[2] Ashton Taylor,[1] Michael Gribskov,[1,†,‡] and Susan S. Taylor[2,‡]

[1]*San Diego Supercomputer Center, University of California San Diego, La Jolla, California*
[2]*Howard Hughes Medical Institute, Department of Chemistry, University of California San Diego, La Jolla, California*

ABSTRACT **The protein kinase superfamily is an important group of enzymes controlling cellular signaling cascades. The increasing amount of available experimental data provides a foundation for deeper understanding of details of signaling systems and the underlying cellular processes. Here, we describe the Protein Kinase Resource, an integrated online service that provides access to information relevant to cell signaling and enables kinase researchers to visualize and analyze the data directly in an online environment. The data set is synchronized with Uniprot and Protein Data Bank (PDB) databases and is regularly updated and verified. Additional annotation includes interactive display of domain composition, cross-references between orthologs and functional mapping to OMIM records. The Protein Kinase Resource provides an integrated view of the protein kinase superfamily by linking data with their visual representation. Thus, human kinases can be mapped onto the human kinome tree via an interactive display. Sequence and structure data can be easily displayed using applications developed for the PKR and integrated with the website and the underlying database. Advanced search mechanisms, such as multiparameter lookup, sequence pattern, and blast search, enable fast access to the desired information, while statistics tools provide the ability to analyze the relationships among the kinases under study. The integration of data presentation and visualization implemented in the Protein Kinase Resource can be adapted by other online providers of scientific data and should become an effective way to access available experimental information. Proteins 2006; 63:78–86.** © **2006 Wiley-Liss, Inc.**

Key words: database; application; visualization; structure; sequence; statistics; kinome

INTRODUCTION

Protein kinases are critical components of cellular signaling cascades that control cell proliferation and responses to external stimuli. Through phosphorylation of their target proteins, they achieve unparalleled levels of control over the finely tuned signaling system of the cell. Malfunction in signaling cascades caused by pathological changes in protein kinase activity has been implicated in a variety of disease conditions, including cancer, inflammation, diabetes, and stroke.[1–7] Protein kinases represent 1.7% of the human genome[8] and account for approximately 4% of plant genomes. Because of the enormous importance of protein kinases for biology and the availability of the large body of structural information that has emerged over the past decade, the protein kinases represent a great opportunity to probe and thoroughly analyze an enzyme family. Because the protein kinases also represent a major class of attractive drug targets, they have received substantial attention from the research community. The structural space of this gene family has filled in over the past decade. The problem of developing drugs targeting protein kinases is twofold: first, the correct target needs to be identified and, second, a specific drug needs to be created in order to modulate activity of the selected kinase with maximal efficacy and minimal side effects. While the second issue lies primarily within the domain of structure-based drug design, chemical biology, structural biology, and molecular biology, the first requires a combined effort by experts in experimental biology, bioinformatics, and computer science. In order to make intelligent decisions regarding composition and functionality of signaling pathways, a significant body of information needs to be generated, analyzed, and stored in an easily accessible form.

To facilitate these efforts, and to provide an integrated view of information describing the structure, function,

Abbreviations: AGC, containing PKA, PKG, PKC families; CAMK, calcium/calmodulin-dependent protein kinases; CKI, casein kinase I; CMGC, containing CDK, MAPK, GSK3, CLK families; OMIM, Online Mendelian Inheritance in Man database; STE, containing yeast Sterile families.

Grant sponsor: National Science Foundation; Grant number: DBI-0217951; Grant sponsor: National Institutes of Health; Grant number: DK-54441.

*These authors have made equal contributions to the reported work.

†M. Gribskov's present address is Department of Biology, Purdue University, West Lafayette, IN 47907.

‡Correspondence to: Michael Gribskov, Department of Biology, Purdue University, West Lafayette, IN 47907. E-mail: gribskov@purdue.edu

‡Correspondence to: Susan S. Taylor, Howard Hughes Medical Center, Department of Chemistry, University of California San Diego, La Jolla, CA 92093. E-mail: staylor@ucsd.edu

Received 2 March 2005; Revised 20 September 2005; Accepted 23 September 2005

Published online 24 January 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20825

sequence, and genetics of protein kinases, the Protein Kinase Resource was established in 1996. It was intended to play the role of a repository of protein kinase–specific information.[9] The original PKR comprised a set of manually curated web pages. Initially, PKR offered visual representations of annotated protein kinase structures, sequence alignments, and protein kinase classification. As PKR further developed, it was converted to a dynamic system based on an underlying relational database to offer a more comprehensive view of the protein kinase superfamily with a better coverage of proteomics and disease states associated with protein kinases.

With the growing amount of experimental data, abundance of structural information, and the need to build stronger links among various types of information, we have created a new version of the Protein Kinase Resource. It implements a novel concept of integration of information access and visualization. In addition, it offers more extensive annotation, coverage of literature, and a variety of software tools for sequence analysis and structure visualization. The new resource complements other efforts in the protein kinase informatics, such as the Alliance for Cell Signaling, which focuses on pathway information and individual molecules involved in signaling. PKR, on the other hand, focuses on the structural aspects, as well as relationships within the protein kinase superfamily, which can be studied through sequence and structure alignments, functional annotations, and evolutionary relationships.

The new Protein Kinase Resource is available at http://www.kinasenet.org.

MATERIALS AND METHODS

Information Content

PKR is a resource dedicated to providing information specific to protein kinases and collates data from several primary resources along with derived information resulting from in-house curation efforts. The database underlying PKR is designed to support the dual nature of its data sources: external resources and internal information. All imported data from external primary resources, such as UniProt, NCBI Taxonomy, and Protein Data Bank (PDB), are maintained in separate schemata that are modeled on the source data. The external data are then linked to protein kinase entries in the internal PKR core database schema. This keeps the curated PKR entries independent of the potentially volatile external data.

Each PKR entry is tied to a unique protein sequence in a distinct organism, and thus corresponds to the basic concept of a UniProt[15] record. PKR draws its core sequence dataset from the UniProt Knowledgebase by using a Java-based parser of the UniProt flatfile. Another major data source is the Protein Data Bank.[16] The simplified set of structural information is parsed from the mmCIF files stored by the PDB. In collaboration with the PDB, we have developed an automatic mechanism to update structural information in PKR to stay current with updated information in PDB and to add new structural information, as it becomes available. This procedure relies on a mapping of accession numbers in primary protein sequence repositories such as UniProt and GenBank that PDB provides for all structures containing polymer entities. The mmCIF dictionary refers to these mappings as “struct_ref”. The update program scans all struct_ref entries referring to UniProt and identifies all structures where the accession number matches a UniProt accession number registered in PKR. The corresponding XML files in the PDB FTP repository are checked for updates based on the file date and subsequently new and updated structures are loaded into PKR.

In addition to data derived directly from the PDB, a subset of protein kinase structures has been aligned with manual curation[14] to produce a high-quality multiple structure alignment that allows for direct comparisons of sequences and structures within the aligned set. At present, there are 16 structures in the manually aligned group. This set is being expanded by automatic alignment to include additional all structures stored in the PKR database.

The multiple sequence alignment was created using the classification of protein kinases in a stepwise approach. First, we created alignments for individual families using ClustalW[11] with limited manual curation. In the next step, the families were combined into groups according to the classification. Extensive manual curation was needed to correct errors that are inevitably introduced when large sets of independently aligned sequences are combined. The resulting alignment contains 4457 sequences as of this writing and provides a reliable view of sequence features for the catalytic domain from the N-terminus to the area in the vicinity of the APE motif at the end of the activation loop (subdomain VIII, according to definition by Hanks and colleagues[13]). The alignment contains a subset of the total content of the database due to the substantial amount of manual curation that had to be used to improve quality. The alignment will continue to expand with more sequences added in the future releases of the resource. For the entries that are not yet represented in the master alignment, the users can run a custom alignment that can align sequences selected either using their classification or displayed lists of search results (as an option in the pulldown menu).

The current protein kinase classification has been derived by clustering of aligned catalytic domain sequences.[10] The resulting classification is similar to the one presented by Manning and colleagues,[8] with the addition of a large group of plant-specific kinases and several differences that primarily concern kinases of unclear lineage that were placed in the middle of the kinome tree by Manning and coworkers.[8] We felt that these kinases should be assigned to specific classes, if possible, because such placement would give a better insight as to the potential functional implications. The group of plant-specific kinases contains mostly transmembrane receptor kinases that share a substantial similarity to each other, but are absent from the human kinome. The entire superfamily was split into nine classes: tyrosine kinases, tyrosine kinase-like (raf-like) kinases, STE-like kinases, CKI kinases, AGC, CMGC,

CAMK, PSK (plant-specific kinases), and MLK (mixed lineage kinases). As of this writing, we have included classification of 5152 entries, which includes the majority of the more intensively studied kinases. We plan to extend the classification to eventually cover the entire content of the database.

In addition to information derived from external sources, we have performed internal curation, which produced domain composition information for kinases that have structural information. The domain content is displayed along with the sequence information and is interactive. We have also annotated orthologous kinases in *Homo sapiens* and several other organisms, such as *Mus musculus*, *C. elegans*, *S. cerevisiae*, and others. This allows for quick comparisons of related kinases across the species, especially in the cases when structural information is available for only some of them. Due to this reason, we have primarily concentrated on entries with structural information. The orthologs can be displayed via a hyperlink in the summary tab of the kinase information display.

Software Architecture

In order to accommodate the diverse types of information hosted by the Protein Kinase Resource, we have used software architecture based on Java programming language. PKR is built upon free open source software: Jakarta Tomcat,[20] Jakarta Struts,[20] and the Hibernate[18,19] object relational persistence framework.

PKR uses the MySQL database management system[23] for information storage, while the Hibernate layer provides the necessary modular structure for data access. PKR is run on an Apache Tomcat web server, which provides an efficient framework for servlets, Java Server Pages, and other elements that deliver dynamic content. All dynamically generated pages in PKR rely on Tiles templates that provide the layout for web pages, but do not contain Java code. The latter feature allows for their maintenance by web designers with no knowledge of programming.

Macromedia Flash™ based inDepth kinase pages were created using Molscript[21] as the molecular visualization engine and POV-Ray[22] to render the scene using ray tracing.

Due to the complex nature of the information provided by the Protein Kinase Resource, we have developed a number of applications that facilitate visualization and analysis of the data. The majority of them are written in the Java programming language that not only provides considerable flexibility at the development stage, but also enables the software to function as a part of the online environment through the use of Java applet and WebStart technology. This approach allowed us to offer the users full-featured molecular visualization applications with capabilities similar to those provided by commercial packages, as well as give the additional ability to work with the database remotely using the PKR Explorer software.

RESULTS AND DISCUSSION

In order to provide the research community with a focused view of the protein kinase superfamily, PKR contains information derived from more general online resources (such as Protein Data Bank, UniProt, or NCBI), as well as internally curated data. This allows the users to obtain information specific to the protein kinase superfamily without searching for this information in sources that have broader coverage. The Protein Kinase Resource provides the researcher with the ability to draw comparisons and assess similarity when presented with either a diverse set of protein kinases or a single protein. At present, the database contains 15,898 protein kinase entries imported from UniProt, and in many cases accompanied by curated information, cross-references to other resources, and literature references. These records are currently selected in collaboration with the UniProt team at the European Bioinformatics Institute by identifying all UniProt entries that contain at least one of the protein sequence features (motifs) associated with protein kinases (Table I).

TABLE I. External Database Feature IDs Used as Criteria for Definition of a Kinase Domain

Feature Type	Feature Id
Interpro	IPR000719
Interpro	IPR002290
Interpro	IPR008271
Pfam	PF00069
Prodom	PD000001
Prosite	PS00107
Prosite	PS00108
Prosite	PS50011
Smart	SM00220

Information provided by PKR can be accessed through a variety of search mechanisms. A quick search facility is present on every page of the web site and returns all entries whose PKR id, name, or description match the entered search term. Advanced search and motif search facilities provide the ability to run more detailed queries. Motif search is used to identify protein kinases that match a particular sequence pattern. Each search returns a list of matching entries with a brief summary of descriptive information — PKR id, name, description, organism, classification, and availability of structural information. For example, running advanced search for human kinases with structures produces 48 hits, which corresponds to the number of unique kinases that may have one or more structures (Fig. 1). Entries from this list can be selected and further analyzed using a pulldown menu to select an action, which is then applied to the selected entries. The actions include displaying the alignment of selected sequences extracted from the pregenerated master alignment (see below), creating a custom alignment using CLUSTALW at the server, running a BLAST search within the PKR sequence database using selected sequence as the query, download of checked sequences as a concatenated FASTA formatted file, opening of the entry in Sequence Viewer or in PKR Explorer or redirection to a detailed inDepth page for the selected protein kinase. The ability to generate alignments of any set of sequences based on search results is especially valuable because it

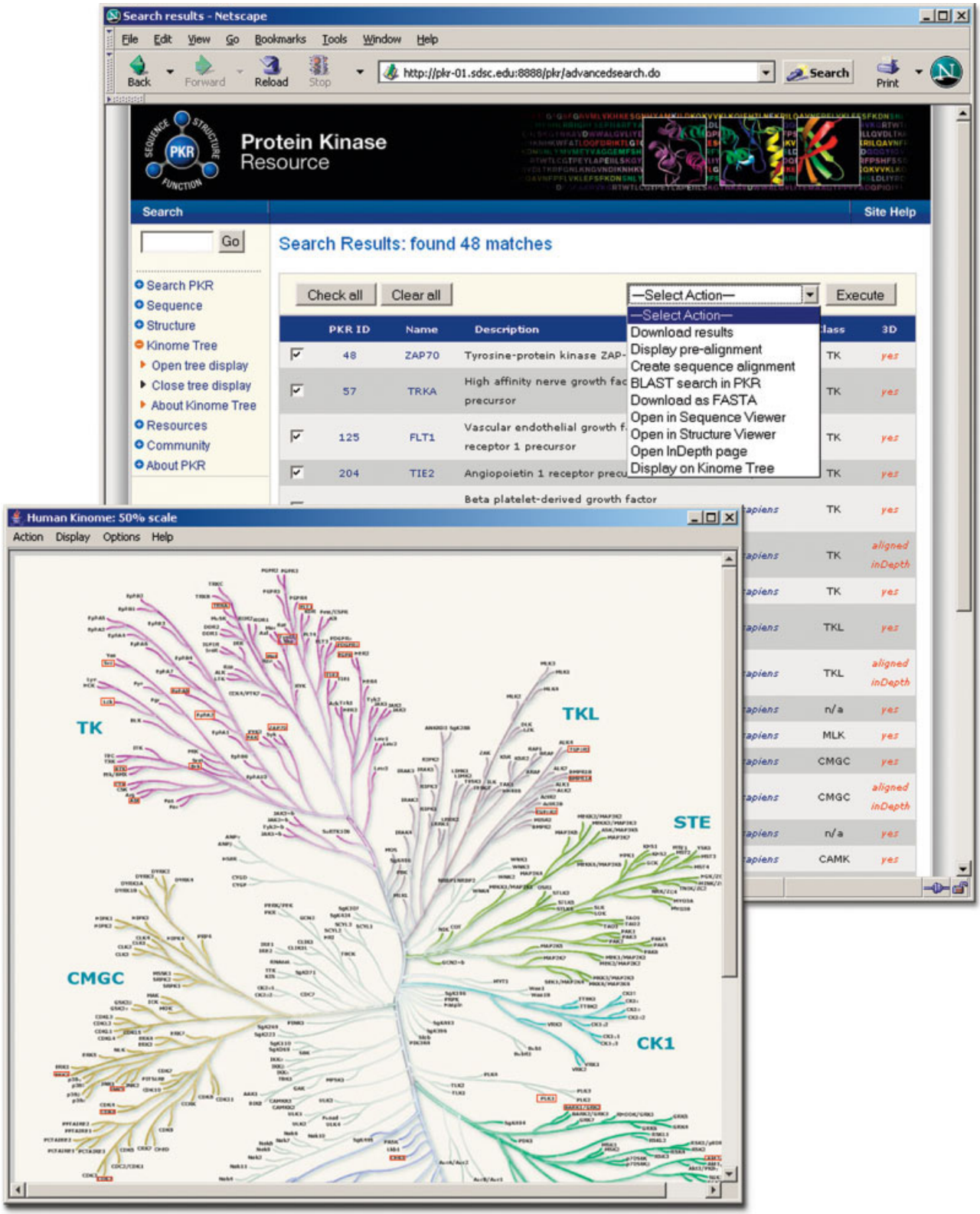

Fig. 1. A list of human kinases that have one or more structures produced by using advanced search. The list can be can be used to display detailed kinase information, create sequence alignments, or download the results. The full range of available options is given in the pulldown menu in the top and bottom panels of the display. The Kinome Tree Viewer is used to map checked entries to the human kinome classification hierarchy. The communication between the main web page and the tree viewer is bidirectional and also permits operations on entries that are selected in the tree, such as alignment, BLAST searches, or basic retrieval of kinase data. The kinase dendrogram image was adapted[8] with permission from Science and Cell Signaling Technology, Inc. (http://www.cellsignal.com). Adapted with permission from Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002;298:1912–1934. Copyright 2002 AAAs.

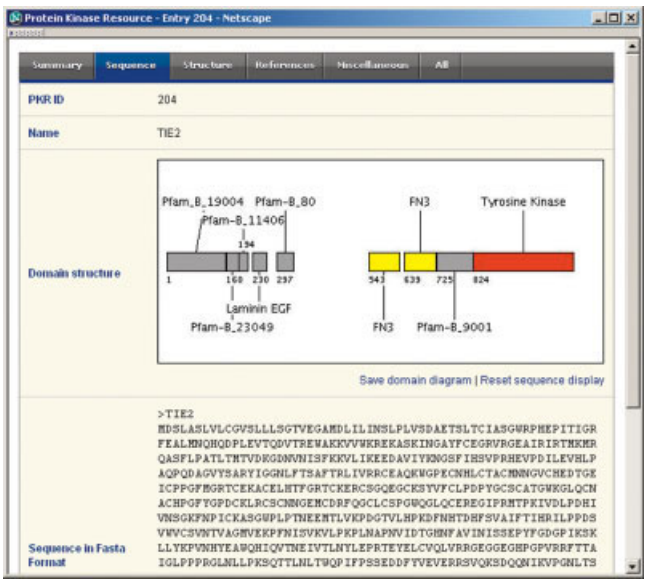

Fig. 2. Annotation information for each kinase is organized in a tabbed interface. The sequence tab contains an interactive domain composition display of the kinase in addition to the amino acid sequence itself. In this case, sequence data for Tie2 protein kinase is displayed. Clicking on the domain cartoon highlights the sequence with the corresponding color, clicking the domain names opens a new window with the description of the domain type at the Pfam web site.

supplements the master alignment and may offer an additional degree of accuracy in the cases of closely homologous groups of kinases. In addition to the listed options, it is possible to map the selected entries to the Kinome Tree Viewer when it is open to look for the position of the kinases of interest on the evolutionary tree.

The entries can be explored in more detail by clicking their names or PKR ids. The information is presented in a tabbed view interface that allows the user to concentrate on the desired type of data. In addition to the annotation, the views feature download links that allow the users to retrieve the sequence or structure files for the entries of interest. The same ability is provided for other types of information, such as sequence alignments and residue conservation diagrams. Annotation includes the name or names, textual description of the protein function, taxonomy, amino acid sequence, isoelectric point, molecular weight, etc. Primary sequences are accompanied by interactive domain composition diagrams for those kinases that have structures. Selection of a domain block highlights the corresponding region of the amino acid sequence providing instant visual feedback (Fig. 2). It is planned to expand this type of annotation to other entries as the work progresses. For entries with known three-dimensional

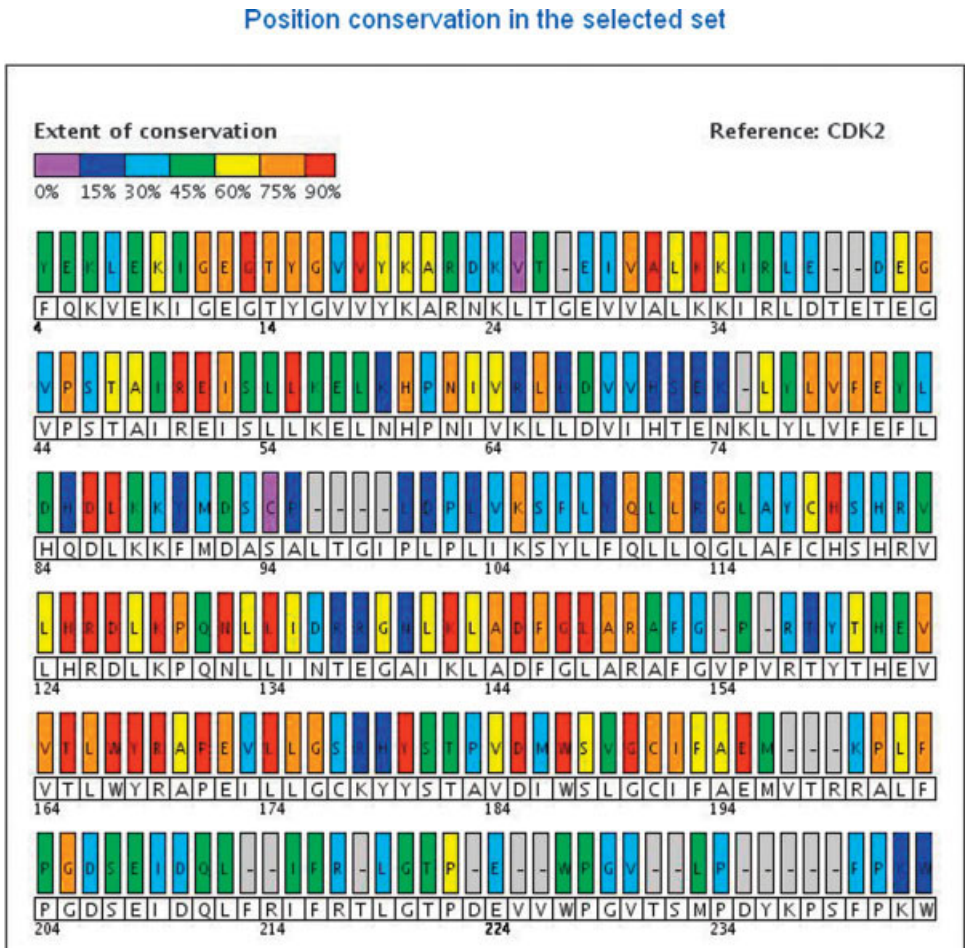

Fig. 3. An example of graphical output of a sequence statistics calculation: side chain conservation within a group of cyclin-dependent kinases using CDK2 as the reference for residue numbering. Coloring of the bars reflects the extent of conservation with warmer colors indicating greater conservation.

structures, PKR provides literature references to structural studies.

One of the main tasks of the new PKR web site is to give a unified view of the superfamily. Inevitably, this translates into the need to provide the means to analyze protein kinases as a group revealing such features as sequence conservation, homology, evolutionary relationships, structural similarities, etc. We have created a master alignment of protein kinase sequences in their catalytic domain. This approach allows users to define sets of sequences, and to evaluate conservation patterns within the group in order to determine differences in primary sequence that could be responsible for observed differences in expression or activity. The results of statistical analyses are presented as color-coded diagrams that show conservation at each position of the alignment or the extent of deviation between the query sequence and the rest of the set (Fig. 3). We have also implemented BLAST[17] search with adjustable parameters. Search results are presented as a list of matching entries sorted in the order of increasing *E*-values (decreasing similarity to the query sequence). This capability enables searching for close homologs of a given query sequence, including those with solved structures.

Because structural information is one of the focal points of PKR, we have implemented structure-specific search that allows retrieval of structural data based on PDB id, description, name of the protein kinase, date range, resolution, organism, and annotation as active or inactive conformation. At present, PKR contains only experimentally determined structures (both the original and aligned coordinates), however, we plan to include models as well. Such theoretical structures accompanied by annotation of the modeling process will be useful in evaluating structural features and/or activity of the given kinase. It has to be noted that the database of the Protein Kinase Resource features structures that contain the catalytic domain of protein kinases, while not focusing on other classes of kinases (e.g., lipid or small molecule kinases).

In the Protein Kinase Resource, we have implemented a novel concept of integration between information delivery and its visualization and analysis. The available data can be retrieved and visualized, further comparisons and analyses can be carried out without the need to find and install additional software. This integration is accomplished via a molecular visualization software platform, PKR Explorer, which serves as a front-end for displaying the protein kinase data. The PKR Explorer is integrated with the new PKR web site to provide context-specific visualization of kinase structures and/or sequence alignments. It is written in Java, which allows its straightforward distribution on the internet and provides a high degree of portability across most major operating systems. The data structures and visualization engine of the PKR Explorer are based on the Molecular Biology Toolkit,[12] developed at San Diego Supercomputer Center. The PKR Explorer has a powerful user interface with an extensive list of visualization features: menu-driven display, solid rendering, multiple coloring options, embedded sequence statistical analysis tools, and various data export options (Fig. 4).

The PKR Explorer has a full set of search and data access capabilities. For example, with this application, it is possible to search for kinases using any of the criteria offered by the web site. In addition, PKR Explorer allows browsing of the database content through an interface in which the visible data fields can be customized or dynamically filtered. PKR Explorer also features a tree viewer for visualization of the protein kinase classification. The tree viewer provides a hierarchical view of evolutionary relationships and allows one to load selected entries into Sequence and Structure Viewers. All statistical analysis tools operate on the loaded set of sequences or on a set selected from the database. Any graphical information resulting from calculations (residue distribution graphs, conservation profiles, etc.) can be saved as images, printed, or exported in formats readable by other applications (such as Microsoft Excel®). At present, the Sequence Viewer component has no built-in alignment capability, and will load only the entries with aligned sequences. If a custom alignment of a specific set of kinases is desired, it should be done through the "Create sequnce alignment" option in the list of kinases, and subsequent opening of the resulting alignment in the Sequence Viewer.

All Java-based applications, such as Kinome Viewer, PKR Explorer, and Sequence Viewer take some time to load when they are used for the first time, up to a minute depending on the network connection. Afterwards, the launch time will be much shorter because the local copy will be used. In addition, the user will be asked whether to trust the content provided by San Diego Supercomputer Center (SDSC), and should answer affirmatively in order to use the applications. In many cases, the dialog window will contain an option for always trusting the content from SDSC. It is recommended to choose it to eliminate future confirmation dialogs when starting the applications.

In order to provide a detailed view of structural features of some important protein kinases, we have established a series of Macromedia Flash® based pages, termed "inDepth pages". This set of interactive pages currently contains structures of 14 protein kinases and shows the subdomain structure in a consistent color-coding scheme. The inDepth pages allow for rapid comparisons to analyze the individual subdomains as first defined by Hanks and Hunter.[13] Initially, the catalytic domain of protein kinase A was divided into 12 subdomains. Using a structure-based alignment by E. Scheeff and P. Bourne,[14] we established subdomain borders. Each subdomain is aligned with and compared to the corresponding subdomain of the other protein kinases. The interactive inDepth pages allow the user to quickly display each subdomain and analyze each residue. Several protein kinase structures are also shown in alignment with other members of the set allowing direct comparison of their folds and conservation of sequence features. Each subdomain can be displayed on a separate page with a detailed representation of its structure and annotation describing the functionally important residues or clusters of residues. This annotation is also imple-

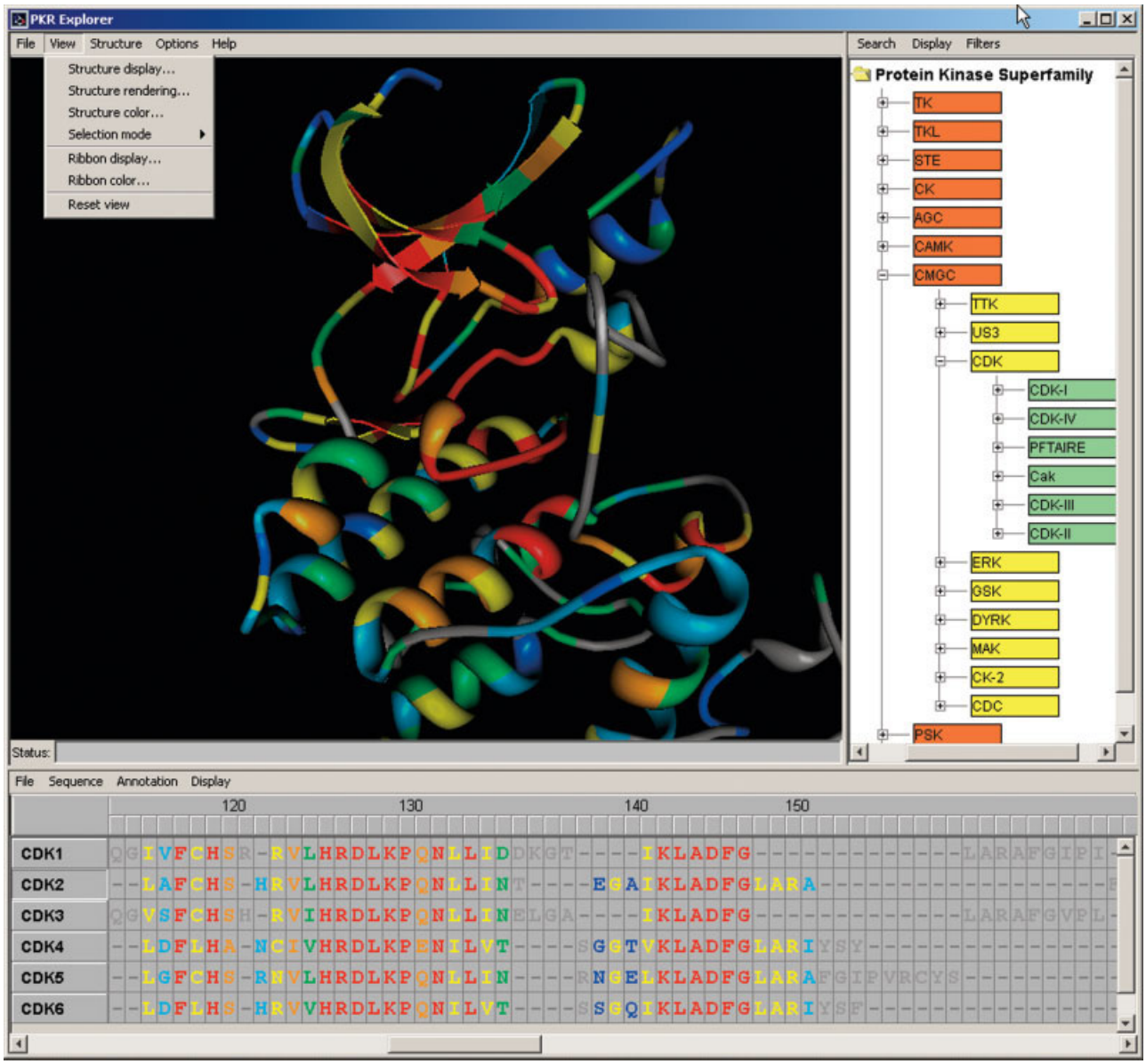

Fig. 4. The PKR Explorer interactive environment. The displayed sequence alignment is for a group of cyclin-dependent kinases and is colored by sequence conservation. The structure is that of CDK2 colored by the same scheme via enabled common coloring mechanism. In this view, it is immediately clear which regions of the structure are more conserved across the family. The classification viewer shows the corresponding part of the hierarchy.

mented in Sequence Viewer, which makes it possible to map the subdomain borders to any protein kinase sequence and structure (through the common coloring mechanism) represented in the master sequence alignment (Fig. 5).

The classification developed for the protein kinases in PKR covers multiple organisms and provides a unified view of the evolutionary relationships between them. However, when only human protein kinases are considered, the classification developed by Manning and colleagues[8] has become the reference point for many researchers ever since its publication. Because the human kinome classification tree is so widely familiar and frequently used, we have incorporated the tree image into PKR as a part of an interactive browsing tool, the Kinome Viewer. It contains the scrollable kinome tree image along with tools that allow one to select the desired protein kinases. As with the other viewers, one can then further analyze or retrieve information on the selected proteins (Fig. 1). At this time, the Kinome Viewer has limited functionality on the MacOS platform due to browser deficiencies. For example, mapping of entries from the web page to the tree is not supported, although the reverse is implemented. On other platforms, including Windows and various types of UNIX, the Kinome Viewer has full functionality with bidirectional communication.

Because the Protein Kinase Resource makes use of several advanced technologies, it will be helpful to visit

Fig. 5. InDepth kinase pages shown displaying a comparison of PKA and CDK2 protein kinases colored by subdomain. The Sequence Explorer displays a set of aligned kinases colored by subdomain.

Site Help page that contains automatic configuration scripts, which determine whether the browser is up to date. If needed, easy-to-follow configuration instructions are given. In addition, illustrated examples of use and available features of the PKR Web site are provided in the online tutorial.

CONCLUSION

The latest version of the Protein Kinase Resource has been designed as an integrated source of information for cell signaling research community. It features a rich set of data, as well as the tools that allow for advanced analysis and various display options. The close integration of data presentation and visualization presents the ability to analyze the data directly within the PKR environment. For example, a novel protein kinase may be compared to a well-studied homolog, their sequences aligned, and the structure of the homologous protein used to identify residues of interest. Three-dimensional structures of proteins can also be directly compared and correlated with sequence alignments using the structure display tools provided by the PKR Explorer. Integration of evolutionary data with the kinome tree provides an additional level of understanding and offers an intuitive way to browse the members of the protein kinase superfamily. The Protein Kinase Resource is a powerful tool that can add another level to our understanding of cell signaling. Integration of annotations, sequence, and structure information can potentially provide a valuable environment for fueling future advances. In addition, the tools and software concepts that we have developed here for the protein kinase superfamily can be easily adapted to any other group of

enzymes. Such interactive tools will be essential if we are to effectively integrate and analyze the wealth of data that is quickly emerging in biology.

ACKNOWLEDGMENTS

The authors thank Dr. Lynn Ten Eyck for his invaluable support of the project and contributions to the development process. We are grateful to members of the Protein Data Bank development team, Wayne Townsend-Merino and Jeff Merino-Ott, for their advice on many aspects of the web application framework we used. In addition, we thank Dr. Philip Bourne and Dr. Eric Scheeff for granting permission to use the manually curated structure alignments of protein kinases. The authors also thank University of California San Diego undergraduate students who helped the project at different stages: Jon Ostrem, Daniel Alyeshmerni, and Nick Le. Jamil Snead, presently at Stanford University, contributed as a summer student.

REFERENCES

1. Bhat RV, Budd Haeberlein SL, Avila J. Glycogen synthase kinase 3: a drug target for CNS therapies. J Neurochem 2004;89:1313–1317.
2. Boehm M, Nabel EG. The cell cycle and cardiovascular diseases. Prog Cell Cycle Res 2003;5:19–30.
3. Cetkovic-Cvrlje M, Tibbles HE. Therapeutic potential of Janus kinase 3 (JAK3) inhibitors. Curr Pharm Des 2004;10:1767–1784.
4. Knudsen BS, Edlund M. Prostate cancer and the met hepatocyte growth factor receptor. Adv Cancer Res 2004;91:31–67.
5. Lin NU, Winer EP. New targets for therapy in breast cancer: small molecule tyrosine kinase inhibitors. Breast Cancer Res 2004;6:204–210.
6. Odabaei G, Chatterjee D, Jazirehi AR, Goodglick L, Yeung K, Bonavida B. Raf-1 kinase inhibitor protein: structure, function, regulation of cell signaling, and pivotal role in apoptosis. Adv Cancer Res 2004;91:169–200.
7. Shah MA, Schwartz GK. Cyclin-dependent kinases as targets for cancer therapy. Cancer Chemother Biol Response Modif 2003;21:145–170.
8. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002;298:1912–1934.
9. Smith CM, Shindyalov IN, Veretnik S, Gribskov M, Taylor SS, Ten Eyck LF, Bourne PE. The protein kinase resource. Trends Biochem Sci 1997;22:444–446.
10. Buzko O, Shokat KM. A comprehensive classification of the protein kinase superfamily. 2005. Manuscript in preparation.
11. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 2003;31:3497–3500.
12. Moreland JL, Gramada A, Buzko O, Zhang Q, Bourne PE. The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. BMC Bioinformatics 2005;6:21.
13. Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science 1988;241:42–52.
14. Scheeff E, Bourne PE. Structural evolution of the protein kinase-like superfamily. PLOS Comp Biol 2005;1(5):e49.
15. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. The Universal Protein Resource (UniProt). Nucleic Acids Res 2005;33:D154–D159.
16. Deshpande N, Addess KJ, Bluhm WF, Merino-Ott JC, Townsend-Merino W, Zhang Q, Knezevich C, Xie L, Chen L, Feng Z, Green RK, Flippen-Anderson JL, Westbrook J, Berman HM, Bourne PE. The RCSB Protein Data Bank: a redesigned query system and relational database based on the mmCIF schema. Nucleic Acids Res 2005;33:D233–D237.
17. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402
18. Hibernate reference documentation. Atlanta, GA: J Boss, Inc.; 2004. Available from http://www.hibernate.org/hib_docs/reference/en/pdf/hibernate_reference.pdf.
19. Bauer C, King G. Hibernate in action, Vol.1. Greenwich, CT: Manning Publications Co.; 2004.
20. Apache Software Foundation. Forest Hill, MD: 2005. http://www.apache.org.
21. Molscript. Stockholm, Sweden: Avatar Software AB; 2005. http://www.avatar.se/molscript/.
22. POV-Ray. Victoria, Austria: Persistence of Vision Raytracer; 2004. Available from: http://www.povray.org.
23. MySQL relational database management system. Uppsala, Sweden: MySQL AB; 2005. http://www.mysql.org.