(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 April 2002 (25.04.2002)

PCT

(10) International Publication Number
WO 02/33596 A2

(51) International Patent Classification[7]: G06F 17/50

(21) International Application Number: PCT/EP01/11955

(22) International Filing Date: 16 October 2001 (16.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00309114.7 17 October 2000 (17.10.2000) EP

(71) Applicant *(for all designated States except US)*: **APPLIED RESEARCH SYSTEMS ARS HOLDING N.V.** [NL/NL]; Pietermaai 15, Curacao (AN).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHURCH, Dennis** [US/CH]; Chemin des Vignes 4, CH-1291 Commugny (CH). **COLINGE, Jacques** [CH/FR]; 8, chemin des Arales, F-74160 Neydens (FR).

(74) Agent: **VOGELSANG-WENKE, Heike**; Grünecker, Kinkeldey, Stockmair & Schwanhäusser, Maximilianstrasse 58, 80538 München (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD OF OPERATING A COMPUTER SYSTEM TO PERFORM A DISCRETE SUBSTRUCTURAL ANALYSIS

(57) Abstract: The invention provides a method of operating a computer system, and a corresponding computer system, for performing a discrete substructural analysis. First, a database of molecular structures is accessed. The database is searchable by molecular structure information and biological and/or chemical properties. In said database, a set of molecules is identified that have a given biological and/or chemical property. Fragments of the molecules in said subset are then determined, and a score value is calculated for each fragment, indicating the contribution of the respective fragment to said given biological and/or chemical property. Finally, a reiteration process is performed by analyzing the determined fragments and calculated scores values, whereby first at least one fragment is selected that has a score value indicating high contribution to said biological and/or chemical property, and then the steps of accessing, identifying, determining and calculating are repeated. Fragments may be any structural subunit of the molecules. The biological and/or chemical properties include biochemical, pharmacological, toxicological, pesticidal, herbicidal and catalytic properties. The invention is preferably used for DNA backsequencing or drug discovery. Preferred embodiments include an reiteration process that increases the fragment size in each iteration, the use of generiy substructures, and an annealing process that glues fragments together.

WO 02/33596 A2