

Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL <i>(Only for new nonprovisional applications under 37 CFR 1.53(b))</i>		Attorney Docket No.	PP6807284
		First Named Inventor	SANGHYUN PARK
		Title	METHOD FOR PREDICTING COMPOUND-PROTEIN BINDING AFFINITY AND APP
		Priority Mail Express® Label No.	
APPLICATION ELEMENTS <i>See MPEP chapter 600 concerning utility patent application contents.</i>		Commissioner for Patents P.O. Box 1450 Alexandria, VA 22313-1450	
<div>1. <input checked="" type="checkbox"/> Fee Transmittal Form (PTO/SB/17 or equivalent)</div> <div>2. <input checked="" type="checkbox"/> Applicant asserts small entity status. See 37 CFR 1.27</div> <div>3. <input type="checkbox"/> Applicant certifies micro entity status. See 37 CFR 1.29. Applicant must attach form PTO/SB/15A or B or equivalent.</div> <div>4. <input checked="" type="checkbox"/> Specification [Total Pages <u>23</u>] Both the claims and abstract must start on a new page. (See MPEP § 608.01(a) for information on the preferred arrangement)</div> <div>5. <input checked="" type="checkbox"/> Drawing(s) (35 U.S.C. 113) [Total Sheets <u>6</u>]</div> <div>6. Inventor's Oath or Declaration [Total Pages <u>4</u>] (including substitute statements under 37 CFR 1.64 and assignments serving as an oath or declaration under 37 CFR 1.63(e))<div>a. <input checked="" type="checkbox"/> Newly executed (original or copy)</div><div>b. <input type="checkbox"/> A copy from a prior application (37 CFR 1.63(d))</div></div> <div>7. <input checked="" type="checkbox"/> Application Data Sheet * See note below. See 37 CFR 1.76 (PTO/AIA/14 or equivalent)</div> <div>8. CD-ROM or CD-R in duplicate, large table, or Computer Program (Appendix)<div><input type="checkbox"/> Landscape Table on CD</div></div> <div>9. Nucleotide and/or Amino Acid Sequence Submission (if applicable, items a. – c. are required)<div>a. <input type="checkbox"/> Computer Readable Form (CRF)</div><div>b. <input type="checkbox"/> Specification Sequence Listing on:<div>i. <input type="checkbox"/> CD-ROM or CD-R (2 copies); or</div><div>ii. <input type="checkbox"/> Paper</div></div><div>c. <input type="checkbox"/> Statements verifying identity of above copies</div></div>			

A Federal agency may not conduct or sponsor, and a person is not required to respond to, nor shall a person be subject to a penalty for failure to comply with an information collection subject to the requirements of the Paperwork Reduction Act of 1995, unless the information collection has a currently valid OMB Control Number. The OMB Control Number for this information collection is 0651-0032. Public burden for this form is estimated to average 12 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the information collection. Send comments regarding this burden estimate or any other aspect of this information collection, including suggestions for reducing this burden to the Chief Administrative Officer, United States Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450 or email InformationCollection@uspto.gov. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. If filing this completed form by mail, send to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. The United States Patent and Trademark Office (USPTO) collects the information in this record under authority of 35 U.S.C. 2. The USPTO's system of records is used to manage all applicant and owner information including name, citizenship, residence, post office address, and other information with respect to inventors and their legal representatives pertaining to the applicant's/owner's activities in connection with the invention for which a patent is sought or has been granted. The applicable Privacy Act System of Records Notice for the information collected in this form is COMMERCE/PAT-TM-7 Patent Application Files, available in the Federal Register at 78 FR 19243 (March 29, 2013). <https://www.govinfo.gov/content/pkg/FR-2013-03-29/pdf/2013-07341.pdf>

Routine uses of the information in this record may include disclosure to: 1) law enforcement, in the event that the system of records indicates a violation or potential violation of law; 2) a Federal, state, local, or international agency, in response to its request; 3) a contractor of the USPTO having need for the information in order to perform a contract; 4) the Department of Justice for determination of whether the Freedom of Information Act (FOIA) requires disclosure of the record; 5) a Member of Congress submitting a request involving an individual to whom the record pertains, when the individual has requested the Member's assistance with respect to the subject matter of the record; 6) a court, magistrate, or administrative tribunal, in the course of presenting evidence, including disclosures to opposing counsel in the course of settlement negotiations; 7) the Administrator, General Services Administration (GSA), or their designee, during an inspection of records conducted by GSA under authority of 44 U.S.C. 2904 and 2906, in accordance with the GSA regulations and any other relevant (i.e., GSA or Commerce) directive, where such disclosure shall not be used to make determinations about individuals; 8) another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)); 9) the Office of Personnel Management (OPM) for personnel research purposes; and 9) the Office of Management and Budget (OMB) for legislative coordination and clearance.

If you do not furnish the information requested on this form, the USPTO may not be able to process and/or examine your submission, which may result in termination of proceedings, abandonment of the application, and/or expiration of the patent.

Additional Uses

Additional USPTO uses of the information in this record may include disclosure to: 1) the International Bureau of the World Intellectual Property Organization, if the record is related to an international application filed under the Patent Cooperation Treaty; 2) the public i) after publication of the application pursuant to 35 U.S.C. 122(b), ii) after issuance of a patent pursuant to 35 U.S.C. 151, iii) if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections, or an issued patent, or iv) without publication of the application or patent under the specific circumstances provided for by 37 CFR 1.14(a)(1)(v)-(vii); and/or 3) the National Archives and Records Administration, for inspection of records.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		
<p>The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.</p> <p>This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.</p>			

Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1					Remove
Legal Name					
Prefix	Given Name	Middle Name	Family Name	Suffix	
	SANGHYUN		PARK		
Residence Information (Select One) <input type="radio"/> US Residency <input checked="" type="radio"/> Non US Residency <input type="radio"/> Active US Military Service					
City	Seoul	Country of Residenceⁱ	KR		
Mailing Address of Inventor:					
Address 1		204-3701, 300 Omok-ro, Yangcheon-gu			
Address 2					
City	Seoul	State/Province			
Postal Code		08004	Countryⁱ	KR	
Inventor 2					Remove
Legal Name					
Prefix	Given Name	Middle Name	Family Name	Suffix	
	SANGMIN		SEO		
Residence Information (Select One) <input type="radio"/> US Residency <input checked="" type="radio"/> Non US Residency <input type="radio"/> Active US Military Service					
City	Jeju-si, Jeju-do	Country of Residenceⁱ	KR		
Mailing Address of Inventor:					
Address 1		#801, 44 Goma-ro			
Address 2					
City	Jeju-si, Jeju-do	State/Province			
Postal Code		63272	Countryⁱ	KR	
Inventor 3					Remove
Legal Name					

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		

Prefix	Given Name	Middle Name	Family Name	Suffix
	SEUNGYEON		CHOI	
Residence Information (Select One) <input type="radio"/> US Residency <input checked="" type="radio"/> Non US Residency <input type="radio"/> Active US Military Service				
City	Seoul	Country of Residence ⁱ	KR	

Mailing Address of Inventor:				
Address 1	101-803, 7-3 Dangsang-ro 45-gil, Yeongdeungpo-gu			
Address 2				
City	Seoul	State/Province		
Postal Code	07213	Country ⁱ	KR	
Inventor 4				Remove
Legal Name				

Prefix	Given Name	Middle Name	Family Name	Suffix
	HWANHEE		KIM	
Residence Information (Select One) <input type="radio"/> US Residency <input checked="" type="radio"/> Non US Residency <input type="radio"/> Active US Military Service				
City	Seoul	Country of Residence ⁱ	KR	

Mailing Address of Inventor:				
Address 1	4F-401, 26 Yonsei-ro 11-gil, Seodaemun-gu			
Address 2				
City	Seoul	State/Province		
Postal Code	03788	Country ⁱ	KR	
Inventor 5				Remove
Legal Name				

Prefix	Given Name	Middle Name	Family Name	Suffix
	JIEUN		LEE	
Residence Information (Select One) <input type="radio"/> US Residency <input checked="" type="radio"/> Non US Residency <input type="radio"/> Active US Military Service				
City	Seoul	Country of Residence ⁱ	KR	

Mailing Address of Inventor:				
Address 1	101-1004, 291 Donggyo-ro, Seodaemun-gu			
Address 2				
City	Seoul	State/Province		
Postal Code	03716	Country ⁱ	KR	

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button.

Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

☐ An Address is being provided for the correspondence information of this application.

Customer Number	183715		
Email Address	bwp@zionip.com	Add Email	Remove Email
Email Address	patent_01@zionip.com	Add Email	Remove Email
Email Address	pbwoong@hotmail.com		Remove Email

Application Information:

Title of the Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		
Attorney Docket Number	PP6807284	Small Entity Status Claimed	<input checked="" type="checkbox"/>
Application Type	Nonprovisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	6	Suggested Figure for Publication (if any)	

Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

Application number of the previously filed application	Filing date (YYYY-MM-DD)	Intellectual Property Authority or Country

Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

35

Representative Information:

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		
<p>Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32). Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.</p>			
Please Select One:	<input checked="" type="radio"/> Customer Number	<input type="radio"/> US Patent Practitioner	<input type="radio"/> Limited Recognition (37 CFR 11.9)
Customer Number	183715		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78. When referring to the current application, please leave the "Application Number" field blank.

Prior Application Status		Remove	
Application Number	Continuity Type	Prior Application Number	Filing or 371(c) Date (YYYY-MM-DD)
Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the Add button.			

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Application Number	Country	Filing Date (YYYY-MM-DD)	Remove
10-2024-0180165	KR	2024-12-06	Access Code ⁱ (if applicable) 50E4
Additional Foreign Priority Data may be generated within this form by selecting the Add button.			

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.

NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number PP6807284	
	Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF	

Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

NOTE: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)

A. Priority Document Exchange (PDX) - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

B. Search Results from U.S. Application to EPO - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

NOTE: Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.			
Applicant 1			
<p>If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.</p>			
<input type="button" value="Clear"/>			
<input checked="" type="radio"/> Assignee	<input type="radio"/> Legal Representative under 35 U.S.C. 117	<input type="radio"/> Joint Inventor	
<input type="radio"/> Person to whom the inventor is obligated to assign.		<input type="radio"/> Person who shows sufficient proprietary interest	
If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:			
Name of the Deceased or Legally Incapacitated Inventor: <input type="text"/>			
If the Applicant is an Organization check here. <input checked="" type="checkbox"/>			
Organization Name	UIF (University Industry Foundation), Yonsei University		
Mailing Address Information For Applicant:			
Address 1	50 Yonsei-ro, Seodaemun-gu		
Address 2			
City	Seoul	State/Province	
Country	KR	Postal Code	03722
Phone Number		Fax Number	
Email Address			
Additional Applicant Data may be generated within this form by selecting the Add button.			

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		

Assignee 1

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

Prefix	Given Name	Middle Name	Family Name	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1			
Address 2			
City		State/Province	
Country ⁱ		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.

Signature:

NOTE: This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

Signature	/Byungwoong Park/		Date (YYYY-MM-DD)	2025-12-06
First Name	Byungwoong	Last Name	Park	Registration Number
80,134				

Additional Signature may be generated within this form by selecting the Add button.

Application Data Sheet 37 CFR 1.76		Attorney Docket Number	PP6807284
		Application Number	
Title of Invention	METHOD FOR PREDICITING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF		

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

- 1 The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
- 2 A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
- 3 A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
- 4 A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
- 5 A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
- 6 A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
- 7 A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
- 8 A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
- 9 A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

METHOD FOR PREDICTING COMPOUND-PROTEIN BINDING AFFINITY AND APPARATUS THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

5 [0001] A claim for priority under 35 U.S.C. § 119 is made to Korean Patent Application No. 10-2024-0180165 filed on November 06, 2024, in the Korean Intellectual Property Office, the entire contents of which are hereby incorporated by reference.

BACKGROUND

10 [0002] Embodiments of the inventive concept described herein relate to a method for predicting compound-protein binding affinity, and an apparatus thereof, and more particularly, relate to a compound-protein binding affinity prediction method that accurately and effectively predicts compound-protein binding affinity based on
15 compound-protein complex free.

[0003] High-throughput screening is used in the early stages of drug development, but it is impossible to evaluate all compound-protein interactions. To compensate for this, protein structure-based virtual screening is employed. For example, molecular docking is widely utilized. However, it has limitations in requiring significant
20 computational resources and having restricted accuracy.

[0004] Moreover, the scarcity of experimentally obtained 3D compound-protein complex structures has limited training datasets, thereby hindering development.

[0005] Accordingly, inventors of the inventive concept endeavored to predict compound-protein binding affinity independently of compound-protein complex
25 structures, thereby culminating in the completion of the inventive concept.

SUMMARY

[0006] Embodiments of the inventive concept provide a compound-protein binding affinity prediction method that accurately and effectively predicts compound-protein
30 binding affinity based on compound-protein complex free.

[0007] Problems to be solved by the inventive concept are not limited to the problems mentioned above, and other problems not mentioned will be clearly understood by those skilled in the art from the following description.

[0008] According to an exemplary embodiment, a compound-protein binding affinity prediction method performed by at least one processor includes receiving compound data and protein data, which interact with each other, generating an attribute vector of a compound and an attribute vector of a protein based on the input compound data and the input protein data, calculating an attention value based on the attribute vector of the compound and the attribute vector of the protein, generating a first interaction matrix based on the attention value, learning a first AI model to predict a binding affinity and a non-covalent interaction of compound-protein by using the first interaction matrix as learning data, and predicting the binding affinity and the non-covalent interaction of the compound-protein based on an output value of the first AI model.

[0009] According to an embodiment of the inventive concept, the receiving of the compound data includes generating a compound structure graph. The compound structure graph has an attribute of an atom as a node and has an attribute of a bond as an edge.

[0010] According to an embodiment of the inventive concept, the calculating of the attention value based on the attribute vector of the compound and the attribute vector of the protein includes calculating a key, a value, and a query based on the attribute vector of the compound and the attribute vector of the protein, and crossing the calculated query of the compound and the calculated query of the protein, or crossing a key and a value of the compound and a key and a value of the protein to provide the crossed result to a sub-attention layer.

[0011] According to an embodiment of the inventive concept, the calculating of the attention value based on the attribute vector of the compound and the attribute vector of the protein further includes providing an output value of the sub-attention layer to a self-attention layer, and providing an output value of the self-attention layer to a feed-forward layer.

[0012] According to an embodiment of the inventive concept, the learning of the first AI model includes calculating an interaction score between one or more elements of the compound and one or more residues of the protein, extracting a latent variable from the score, and providing the latent variable to a fully connected layer.

[0013] According to an embodiment of the inventive concept, the output value of the first AI model includes the attribute vector of the compound, the attribute vector of

the protein, the first interaction matrix, and the binding affinity predicted by the first AI model.

[0014] According to an embodiment of the inventive concept, the predicting of the binding affinity and the non-covalent interaction of the compound-protein based on the output value of the first AI model includes performing knowledge distillation by inputting the output value of the first AI model into a second AI model, and predicting the non-covalent interaction and the binding affinity in the second AI model. Data received by the first AI model is based on a compound-protein complex.

[0015] According to an embodiment of the inventive concept, the performing of the knowledge distillation includes calculating a loss function by comparing the first interaction matrix with a second interaction matrix generated by the second AI model.

[0016] According to an embodiment of the inventive concept, a computer program stored in a computer-readable recording medium is provided to execute the above-described method on a computer.

[0017] According to an exemplary embodiment, a computing apparatus includes a communication module, a memory, and at least one processor connected to the memory and configured to execute at least one computer-readable program included in the memory. The at least one program includes instructions for receiving a structure of a compound and a structure of a protein, which interact with each other, generating an attribute vector of a compound and an attribute vector of a protein based on the input structure of the compound and the input structure of the protein, calculating an attention value based on the attribute vector of the compound and the attribute vector of the protein, generating a first interaction matrix based on the attention value, learning a first AI model to predict a binding affinity and a non-covalent interaction of compound-protein by using the first interaction matrix as learning data, and predicting the binding affinity and the non-covalent interaction of the compound-protein based on an output value of the first AI model.

BRIEF DESCRIPTION OF THE FIGURES

[0018] The above and other objects and features will become apparent from the following description with reference to the following figures, wherein like reference numerals refer to like parts throughout the various figures unless otherwise specified,

and wherein:

[0019] FIG. 1 is a functional block diagram illustrating an internal configuration of a computing apparatus, according to an embodiment;

5 [0020] FIG. 2 is a functional block diagram showing an internal configuration of an AI model, according to an embodiment;

[0021] FIG. 3 is a flowchart of a compound-protein binding affinity prediction method, according to an embodiment;

[0022] FIG. 4 is a diagram illustrating an operation of calculating an attention value by using a layer, according to an embodiment;

10 [0023] FIG. 5 is a graph for comparing the performance of models predicting a binding affinity of compound-protein, according to an embodiment; and

[0024] FIG. 6 is a block diagram showing a hardware configuration of a computing apparatus, according to an embodiment.

15 DETAILED DESCRIPTION

[0025] Hereinafter, embodiments of the inventive concept will be described in detail with reference to the accompanying drawings. However, the inventive concept is not intended to be limited or restricted by embodiments. Unless otherwise defined, all terms (including technical and scientific terms) used in the specification should have
20 the same meaning as commonly understood by those skilled in the art to which the inventive concept pertains, but which may vary according to the intent or precedent of those practicing in the art, the emergence of new technology, and the like.

[0026] Moreover, terms, such as those defined in commonly used dictionaries, should not be interpreted in an idealized or overly formal sense unless expressly so
25 defined herein. Terms arbitrarily selected by the applicant of embodiments may also be used in a specific case. In this case, the detailed meanings are given in the corresponding description. Hence, these terms used in the inventive concept may be defined based on their meanings and the contents of the inventive concept, not by simply stating the terms.

30 [0027] It will be understood that the terms “comprises,” “comprising,” “includes” and/or “including,” when used herein, specify the presence of stated elements and/or components, but do not preclude the presence or addition of one or more other elements and/or components. Moreover, as used in the specification, the singular

terms “a,” “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. Besides, the expression “at least one of a, b, and/or c” described throughout this specification may encompass ‘a alone’, ‘b alone’, ‘c alone’, ‘a and b’, ‘a and c’, ‘b and c’, or ‘all of a, b, and c’.

5 [0028] In the meantime, the term “first and/or second” used in the specification will be used to describe various elements but will be described only for the purpose of distinguishing one element from another element, not limiting an element of the corresponding term. For example, without departing the scope of the inventive concept, a first element may be referred to as a second element, and similarly, a
10 second element may be referred to as a first element.

[0029] In addition, terms such as “...unit”, “...module”, etc. described in this specification mean a unit that processes at least one function or operation, which may be implemented as hardware or software, or a combination of hardware and software. Furthermore, an embodiment of the inventive concept described herein may be
15 represented by functional block configurations and various processing steps. These functional blocks may be implemented in the variable number of hardware or/and software configurations that perform specific functions. For example, embodiments of the inventive concept may employ integrated circuit configurations such as a memory, processing, logic, a look-up table, etc., which may execute various
20 functions under the control of one or more microprocessors or other control devices.

[0030] In an embodiment of the inventive concept, functions related to artificial intelligence may be implemented through a processor and a memory. In this case, the processor may be one of a general-purpose processor such as a center processing unit (CPU), an application processor (AP), a digital signal processor (DSP), a graphics-
25 dedicated processor such as a graphic processing unit (GPU), a vision processing unit (VPU), and an AI-dedicated processor such as a neural network processing unit (NPU). The processor may process input data depending on an AI model or a predefined operating rule, which is stored in the memory. Alternatively, when the processor is an AI-dedicated processor, the AI-dedicated processor may be designed
30 with a hardware structure specialized for the processing of a specific AI model. In some embodiments of the inventive concept, functions related to artificial intelligence may be implemented through a plurality of processors.

[0031] In an embodiment of the inventive concept, the predefined operating rule or

the artificial intelligence model may be configured to perform machine learning. Here, being configured to perform machine learning means that the predefined operation rule or the artificial intelligence model is configured to perform a desired feature (or purpose) by learning pieces of learning data based on a learning
5 algorithm. This learning may be performed by a device itself, on which the artificial intelligence according to an embodiment of the inventive concept is implemented, or may be performed through a separate server and/or system.

[0032] The artificial intelligence model may be implemented with a neural network (or an artificial neural network) and may operate based on a statistical learning
10 algorithm that mimics biological neurons in machine learning and cognitive science. The neural network may refer to a model as a whole having the ability to solve problems as artificial neurons (nodes), which form a network by connecting synapses, changes the strength of their synaptic connections through learning. The neural network may be composed of a plurality of neural network layers. For
15 example, the neural network may include an input layer, a hidden layer, and an output layer. Each of the plurality of neural network layers may include at least one node and at least one weight, and may perform neural network operations through operations between weights and the operation results of the previous layer. At least
20 one weight of the plurality neural network layers may be optimized by the training result of the artificial intelligence model. For example, during the training process, the at least one weight may be updated such that a loss value or cost value obtained from the artificial intelligence model is reduced or minimized. The neural network may infer the desired result from an arbitrary input.

[0033] Training methods of the artificial intelligence model may be classified into
25 supervised learning, in which input data and output data are provided as training data according to the learning method, and the correct answer (output data) corresponding to the problem (input data) is determined, unsupervised learning, in which only the input data is provided without the output data, and the correct answer (output data) corresponding to the problem (input data) is not determined, reinforcement learning,
30 in which a reward is given whenever an action is taken in a current state, and training proceeds to maximize the reward, and the like. Alternatively, the training methods may be distinguished based on the architecture, which is the structure of the learning model.

[0034] According to an embodiment of the inventive concept, the artificial intelligence model may use at least one of various artificial intelligence structures and algorithms such as a convolution neural network (CNN) (e.g., GoogleNet, AlexNet, or VGG Network), a region with convolution neural network (R-CNN), a region proposal network (RPN), a recurrent neural network (RNN), a stacking-based deep neural network (S-DNN), a state-space dynamic neural network (S-SDNN), a deconvolution network, a deep belief network (DBN), a restricted Boltzman machine (RBM), a fully convolutional network, a long short-term memory (LSTM) Network, a classification network, Generative Modeling, eXplainable AI, Continual AI, Representation Learning, AI for Material Design, algorithms for natural language processing (e.g., BERT, SP-BERT, MRC/QA, Text Analysis, Dialog System, GPT-3, and GPT-4), algorithms for vision processing (e.g., Visual Analytics, Visual Understanding, Video Synthesis, and ResNet), algorithms for data intelligence (e.g., Anomaly Detection, Prediction, Time-Series Forecasting, Optimization, Recommendation, and Data Creation), but is not limited thereto. The above-described examples are merely illustrative of artificial intelligence structures and algorithms used in accordance with embodiments of the inventive concept, and do not limit the artificial intelligence structures and algorithms used in accordance with embodiments of the inventive concept.

[0035] Hereinafter, various embodiments of the inventive concept will be described in detail with reference to the accompanying drawings. In describing an embodiment, technical details that are well known in the art to which the inventive concept pertains and are not directly related to the inventive concept will be omitted. This is to avoid obscuring the essence of the inventive concept and to convey it more clearly by omitting unnecessary explanations. For the same reason, some components in the attached drawings are exaggerated, omitted, or shown schematically. Furthermore, the size of each component does not necessarily reflect its actual size. In this specification, the same reference numerals throughout the specification may refer to the same or corresponding components.

[0036] In this specification, the term ‘interaction’ may refer to a non-covalent interaction between a compound and a protein, specifically a non-covalent interaction between elements of the compound and residues of the protein.

[0037] In this specification, the term ‘binding affinity’ may refer to the strength of

the binding interaction between a compound and a protein.

[0038] In this specification, the term ‘knowledge distillation’ may refer to a method by which a teacher model transfers learned knowledge to a student model. In this specification, the teacher model may be a first Artificial intelligence (AI) model 100,
5 and the student model may be a second AI model.

[0039] FIG. 1 is a functional block diagram illustrating an internal configuration of a computing apparatus 1000, according to an embodiment. Referring to FIG. 1, the computing apparatus 1000 may refer to any AI device that receives protein data 101 and compound data 102 to predict the binding affinity and the non-covalent binding
10 of the corresponding protein-compound. The computing apparatus 1000 may include a first AI model 100 and a second AI model 200, and the first AI model 100 and the second AI model 200 may receive the protein data 101 and the compound data 102.

[0040] The first AI model 100 and the second AI model 200 are described as physically separate components from each other. However, this is only an example.
15 In another embodiment, the first AI model 100 and the second AI model 200 may be logically separate structures, in which case they may be implemented by separate functions on a single server.

[0041] The first AI model 100 may be a device learned to predict binding affinity and non-covalent interactions of compound-protein by using protein data and compound data. The first AI model 100 may perform knowledge distillation by
20 delivering an output value to the second AI model 200.

[0042] The second AI model 200 may predict the binding affinity and the non-covalent interaction of compound-protein based on the protein data, the compound data, and data received from the first AI model 100.

25 [0043] A loss function may be calculated based on a first output value (a long dashed-dotted-dotted line 1) of the first AI model 100 and the first output value (a long dashed-dotted-dotted line 2) of the second AI model 200. Moreover, the non-covalent interaction and the binding affinity may be predicted based on a second output value (a dashed-dotted line 3) of the first AI model 100 and a second output
30 value (a dashed-dotted line 4) of the second AI model 200. The method for predicting the binding affinity and the non-covalent interaction of compound-protein by using the first AI model 100 and the second AI model 200 of the computing apparatus 1000 will be described in detail below.

[0044] FIG. 2 is a functional block diagram showing an internal configuration of an AI model, according to an embodiment.

5 [0045] According to an embodiment, both the first AI model 100 and the second AI model 200 may include a protein encoding module 110, a compound encoding module 120, an attention module 130, and a prediction module 140. Hereinafter, the descriptions of the protein encoding module 110, the compound encoding module 120, the attention module 130, and the prediction module 140 may be applied to both the first AI model 100 and the second AI model 200.

10 [0046] The first AI model 100 and the second AI model 200 may receive the protein data 101 and the compound data 102.

[0047] The first AI model 100 may obtain the protein data 101 and the compound data 102 from a three-dimensional (3D) structure of a compound-protein complex, for example, from a PDBbind database.

15 [0048] The second AI model 200 may obtain the protein data 101 and the compound data 102 from a free compound-protein complex model, for example, from a BindingDB database.

[0049] Here, the protein data 101 may include a ligand of a pocket region, which is a recessed region on the protein surface, and the compound data 102 may include a target molecule region, a region that binds to the protein.

20 [0050] When the protein data 101 and the compound data 102 are received, the protein encoding module 110 and the compound encoding module 120 may perform encoding to use the protein data 101 and the compound data 102 as an input of the attention module 130. Here, the encoding may refer to a method of converting original data into a different format.

25 [0051] The protein encoding module 110 and the compound encoding module 120 may respectively generate a protein attribute vector and a compound attribute vector based on the protein data 101 and the compound data 102, and may transmit the protein attribute vector and the compound attribute vector to the attention module 130. In detail, the protein encoding module 110 and the compound encoding module 30 120 may obtain residue information of a protein pocket and element information of a compound by performing embedding based on the protein data 101 and the compound data 102.

[0052] Here, the “embedding” may refer to converting a sequence (e.g., an amino

acid sequence or an element of a compound) into a vector having a specific numerical value to quantify data.

5 [0053] When receiving a protein attribute vector and a compound attribute vector, the attention module 130 may calculate an attention value based on the protein attribute vector and the compound attribute vector. Here, the attention value may represent the interaction between the compound and the protein. When receiving the protein attribute vector, the attention module 130 may calculate a protein attention value and transmit the protein attention value to the prediction module 140. When receiving the compound attribute vector, the attention module 130 may calculate a
10 compound attention value and transmit the compound attention value to the prediction module 140.

[0054] When receiving the protein attention value and the compound attention value, the prediction module 140 may generate an interaction matrix based on the protein attention value and the compound attention value. In detail, the prediction module
15 140 of the first AI model 100 may generate a first interaction matrix, and the prediction module 140 of the second AI model 200 may generate a second interaction matrix. Here, the interaction matrix may be a matrix representing an interaction score between a residue of a protein and an element of a compound.

[0055] Referring to FIGS. 1 and 2, the first AI model 100 may perform learning to
20 predict the binding affinity and the non-covalent interaction of a compound-protein complex by using the first interaction matrix as learning data. The first AI model 100 may receive the compound-protein complex and may transmit the corresponding a binding affinity prediction value and a non-covalent interaction prediction value of the compound-protein complex to the second AI model 200.

25 [0056] The second AI model 200 may learn the binding affinity and the non-covalent interaction of the compound-protein complex based on the output value of the first AI model and the second interaction matrix.

[0057] In this case, the second AI model 200 may be based on complex free, but the second AI model 200 may accurately predict the binding affinity of compound-
30 protein by receiving knowledge distillation on complex-based data from the first AI model 100.

[0058] The specific details of a compound-protein binding affinity prediction method according to an embodiment will be described in detail in FIGS. 3 and 4 below.

[0059] FIG. 3 is a flowchart of a compound-protein binding affinity prediction method, according to an embodiment. Referring to FIG. 3, the binding affinity prediction method may include an operation of inputting compound data and protein data (S100), an operation of generating a compound attribute vector and a protein attribute vector (S200), an operation of calculating an attention value based on a vector (S300), an operation of generating a first interaction matrix and learning a first AI model (S400), an operation of generating a second interaction matrix and learning a second AI model (S500), and an operation of making a prediction in the second AI model (S600). Hereinafter, operations S100 to S300 may be performed by both the first AI model 100 and the second AI model 200.

[0060] Operation S100 may refer to a step in which the encoding modules 110 and 120 of the AI models 100 and 200 receive compound data and protein data. Here, the compound data and the protein data may be provided as compound-protein complex data in the first AI model 100 and may be provided as compound-protein complex free data in the second AI model 200.

[0061] The received protein data may include amino acid sequences of a residues in a pocket region that bind to a compound. Here, the second AI model 200 may use a protein binding pocket prediction tool to extract the input amino acid sequences associated with a binding pocket of the protein by receiving the compound-protein complex free data. Afterwards, the second AI model 200 may extract residue amino acid embedding of the pocket region by using another pre-learned AI model.

[0062] In an embodiment, an operation of receiving the compound data may include an operation of generating a compound graph. The compound graph may be generated using a known AI graph. Here, the compound structure graph may have atomic attributes as nodes and bond attributes as edges.

[0063] In this case, the compound data may include an elemental attribute and a bond attribute. For example, the element attribute may include an element's type, chirality, formal charge, the number of bonded hydrogen atoms, the number of bonded free electrons, hybridization, aromaticity, cyclicity, or the like. Here, the bond may refer to a bond between elements of a compound. For example, the bond attribute may include the type of bond (a single bond, a double bond, or a triple bond) and the nature of bond (directionality, chirality, stereo, or conjugation).

[0064] Operation S200 may refer to a step for generating a compound attribute

vector and a protein attribute vector in the encoding modules 110 and 120 of the AI models 100 and 200.

[0065] In detail, the protein encoding module 110 may extract the protein attribute vector by using the amino acid embedding of a residue as an input. For example, the protein attribute vector may be extracted by using the amino acid embedding of the residue as an input based on Equation 1 below.

[Equation 1]

$$p_j = \text{SelfAttLayer}_{P \rightarrow P} (x_j, \{x_1, \dots, x_{L_p}\})$$

[0066] Here, p_j may denote the extracted protein attribute vector, and x_j may denote the amino acid of a residue.

[0067] The compound encoding module 120 may extract a compound attribute vector by using the node and edge of a compound structure graph as inputs. For example, the compound attribute vector may be extracted by using the node and edge of a compound structure graph as inputs based on Equation 2 below.

[Equation 2]

$$h_v^{(l)} = \text{COMBINE}(h_v^{(l-1)}, \text{AGGREGATE}(\{(h_v^{(l-1)}, h_u^{(l-1)}, e_{uv}) : u \in \mathcal{N}(v)\}))$$

[0068] Here, $\mathcal{N}(v)$ may denote all neighbors of node v ; e_{uv} may denote the edge between nodes u and v ; and, $h_v^{(l-1)}$ may denote the $(l-1)$ -th layer of node v .

[0069] This layer may be repeatedly applied to update nodes, and the updated nodes may be used as the compound attribute vector.

[0070] Operation S300 may be a step in which the attention module 130 of the AI models 100 and 200 calculates an attention value based on the compound attribute vector and the protein attribute vector.

[0071] In particular, the attention module 130 may calculate each key, each value, and each query based on the compound attribute vector and the protein attribute vector. Here, each key, each value, and each query may be calculated by using a known attention algorithm.

[0072] The attention module 130 may include a cross-attention layer to perform cross-attention. Accordingly, the attention module 130 may cross the calculated compound query with the calculated protein query, or cross the key and value of a compound with the key and value of a protein, and may provide the crossed result to a sub-attention layer.

[0073] The sub-attention layer may perform attention by using a multi-head self-attention layer and may provide an output value to a self-attention layer. The self-attention layer may perform self-attention to provide an output value to a feed-forward layer. The feed-forward layer may perform a feed-forward operation to provide a compound attention value and a protein attention value. Accordingly, when the attention module 130 receives a compound attribute vector, the output value may be the compound attention value. When the attention module 130 receives a protein attribute vector, the output value may be the protein attention value.

[0074] FIG. 4 is a diagram illustrating an operation of calculating an attention value by using a layer, according to an embodiment. Referring to FIG. 4, the attention module 130 may cross a query of a compound attribute vector 'v' with the query of a protein attribute vector 'p', or may cross the key and value of the compound attribute vector 'v' with the key and value of the protein attribute vector 'p'. The attention module 130 may include a sub-attention layer, a self-attention layer, and a feed-forward layer to calculate a compound attention value and a protein attention value by using multi-head attention, self-attention, and feed-forward. The finally calculated compound attention value may be represented as \bar{v}_i , and the finally calculated protein attention value may be represented as \bar{p}_j .

[0075] Operation S400 may be a step in which the prediction module 140 of the first AI model 100 generates a first interaction matrix and learns a first AI model. Here, the first interaction matrix may represent an interaction score between the element of a compound and the residue of a protein. The interaction score may be calculated based on Equation 3 below.

[Equation 3]

$$M_{ij} = \sigma(\bar{v}_i W_1) \cdot \sigma(\bar{p}_j W_2)^T$$

[0076] Here, M_{ij} represents the interaction score between the i-th element and the j-th residue, and W_1 and W_2 represent learning parameters, respectively.

[0077] The prediction module 140 of the first AI model 100 may extract a latent variable based on the generated interaction score.

[Equation 4]

$$V_{inter} = \sum_{i=1}^{N_a} \sum_{j=1}^{L_p} M_{ij} [\sigma(\bar{v}_i W_3), \sigma(\bar{p}_j W_4)]$$

[0078] Here, V_{inter} represents the latent variable, and W_3 and W_4 represent learning parameters, respectively.

[0079] The prediction module 140 of the first AI model 100 may provide the latent variable to a fully connected layer to predict a binding affinity.

[0080] Here, a loss function of the first AI model 100 may be expressed based on Equation 5 below.

[Equation 5]

$$\mathcal{L}_{total} = \mathcal{L}_{aff} + \alpha_T \mathcal{L}_{inter}$$

[0081] Here, α_T denotes a weight parameter, and \mathcal{L}_{aff} and \mathcal{L}_{inter} may be expressed by Equations 6 and 7, respectively.

[Equation 6]

$$\mathcal{L}_{aff} = \frac{1}{N} \sum_{n=1}^N \|P_{gt} - P_{pred}\|_n^2$$

[Equation 7]

$$\mathcal{L}_{inter} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{N_a} \sum_{j=1}^{L_p} -\|B_{ij} \log M_{ij} + (1 - B_{ij}) \log (1 - M_{ij})\|_n$$

[0082] Here, P_{gt} may denote an actual binding affinity; P_{pred} may denote a predicted binding affinity; and B_{ij} may denote a binary label between an actual i-th element and a j-th residue.

[0083] The first AI model 100 may deliver the output value of the first AI model to the second AI model 200. Here, the output value of the first AI model may include the predicted binding affinity, the attribute vector of a compound, the attribute vector of a protein, and a first interaction matrix.

[0084] Operation S500 may be a step in which the prediction module 140 of the second AI model 200 generates a second interaction matrix and learns a second AI model. In detail, operation S500 may refer to a step of inputting the output value of

the first AI model into the second AI model to perform knowledge distillation, and may be a step in which the prediction module 140 of the second AI model 200 generates the second interaction matrix and learns the second AI model based on the output value of the first AI model. Here, the output value of the first AI model may include the predicted binding affinity, the attribute vector of a compound, the attribute vector of a protein, and a first interaction matrix.

[0085] The prediction module 140 of the second AI model 200 may learn the second AI model by generating the second interaction matrix by using the same procedure as operations S100 to S400, except that it receives compound-protein complex free data.

[0086] Operation S600 may refer to a step of predicting a non-covalent binding interaction and a binding affinity from the learned second AI model 200. In an embodiment, operation S600 may include an operation of comparing the first interaction matrix and the second interaction matrix to derive a loss function.

[0087] In particular, the second AI model 200 may calculate two loss functions through a two-stage optimization process. The first loss function may be expressed by Equations 8 and 9 below.

[Equation 8]

$$\mathcal{L}_{stage_1} = \mathcal{L}_{inter} + \mathcal{L}_{hint}$$

[Equation 9]

$$\begin{aligned} \mathcal{L}_{inter} &= \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^{N_a} \sum_{j=1}^{L_p} -\|\widehat{M}_{ij}^T \log M_{ij}^S + (1 - \widehat{M}_{ij}^T) \log (1 - M_{ij}^S)\|_n \\ \mathcal{L}_{hint} &= 0.5 \left(\frac{1}{N} \sum_{n=1}^N \Phi_n \|\psi_T^c(I; W_{hint}^c) - \psi_S^c(I; W_{guide}^c)\|_n^2 + \right. \\ &\quad \left. \frac{1}{N} \sum_{n=1}^N \Phi_n \|\psi_T^p(I; W_{hint}^p) - \psi_S^p(I; W_{guide}^p)\|_n^2 \right) \end{aligned}$$

[0088] Here, M_{ij}^S may denote the interaction between the i-th element and the j-th residue predicted by the second AI model 200; \widehat{M}_{ij}^T may denote the interaction between the i-th element and the j-th residue predicted by the first AI model 100;

W_{hint} and W_{guide} may denote a parameter of a layer for calculating an attention value based on a compound attribute vector, and a parameter of a layer for calculating an attention value based on a protein attribute vector, respectively; ψ_T and ψ_S may denote deep neural functions of the first AI model 100 and the second AI model 200, respectively; and, Φ_n may denote a parameter that quantizes the confidence for the prediction of the first AI model 100.

[0089] The second loss function calculated from the second AI model 200 may be expressed based on Equation 10 below.

[Equation 10]

$$\mathcal{L}_{stage_2} = \mathcal{L}_{reg} + \alpha_S \mathcal{L}_{imit}$$

$$\mathcal{L}_{reg} = \frac{1}{N} \sum_{n=1}^N \|P_S - P_{gt}\|_n^2$$

$$\mathcal{L}_{imit} = \frac{1}{N} \sum_{n=1}^N \Phi_n \|P_S - P_T\|_n^2$$

[0090] Here, α_S may denote a weight parameter controlling the final predicted affinity, and P_S , P_T , and P_{gt} may denote the prediction of the second AI model 200, the prediction of the first AI model 100, and an actual label, respectively.

[0091] FIG. 5 is a graph for comparing the performance of models predicting a binding affinity of compound-protein, according to an embodiment. In FIG. 5, Blendnet(s) represents the second AI model 200 according to an embodiment of the inventive concept. Referring to FIG. 5, compared to other prediction models, the prediction method according to an embodiment of the inventive concept demonstrates superior prediction performance in all of new protein-based segmentation, new compound-based segmentation, and random segmentation model evaluation.

[0092] FIG. 6 is a block diagram showing a hardware configuration of a computing apparatus, according to an embodiment.

[0093] The computing apparatus 1000 may include a memory 1100, a processor 1200, a communication module 1300, and an input/output interface 1400. As shown

in FIG. 6, the computing apparatus 1000 may be configured to exchange information and/or data over a network by using the communication module 1300.

[0094] The memory 1100 may include any computer-readable recording medium. According to an embodiment, the memory 1100 may include a permanent mass storage device such as a random access memory (RAM), a read only memory (ROM), a disk drive, a solid state drive (SSD), a flash memory, or the like. For another example, the permanent mass storage device such as a ROM, a SSD, a flash memory, or a disk drive may be included in the computing apparatus 1000 as a permanent storage device separate from the memory. Moreover, the memory 1100 may store the first AI model 100 and the second AI model 200, and may store an operating system and at least one program code.

[0095] These software components may be loaded from a computer-readable recording medium independent of the memory 1100. Such the separate computer-readable recording medium may include a recording medium capable of being directly connected to the computing apparatus 1000, and may include, for example, a computer-readable recording medium such as a floppy drive, a disk, a tape, a DVD/CD-ROM drive, and a memory card. For another example, the software components may be loaded into the memory 1100 through the communication module 1300, not the computer-readable recording medium. For example, at least one program may be loaded into the memory 1100 based on a computer program installed by files provided by developers or a file distribution system, which distributes a file for installing an application, through the communication module 1300.

[0096] The processor 1200 may be configured to process instructions of a computer program by performing basic arithmetic, logic, and input and output operations. The instructions may be provided to another user terminal (not shown) or another external system by the memory 1100 or the communication module 1300.

[0097] The communication module 1300 may provide a configuration or function that allows a user terminal (not shown) and the computing apparatus 1000 to communicate with each other over the network. The computing apparatus 1000 may provide a configuration or function for communicating with an external system (e.g., a separate cloud system, etc.). For example, control signals, commands, data, or the like provided under the control of the processor 1200 of the computing apparatus

1000 may be transmitted to a user terminal and/or the external system through the communication module of the user terminal and/or the external system via the communication module 1300 and a network.

5 [0098] Moreover, the input/output interface 1400 of the computing apparatus 1000 may be a means for interfacing with an apparatus (not shown) for an input or an output, which is connected to the computing apparatus 1000 or is included in the computing apparatus 1000. In FIG. 6, the input/output interface 1400 is shown as an element configured separately from the processor 1200, but is not limited thereto. For example, the input/output interface 1400 may be configured to be included in the
10 processor 1200. The computing apparatus 1000 may include more components than those of FIG. 6. However, there is no need to clearly illustrate most conventional components.

[0099] The processor 1200 of the computing apparatus 1000 may be configured to manage, process, and/or store information and/or data received from a plurality of
15 user terminals and/or a plurality of external systems.

[00100] The above-described method and/or various embodiments may be implemented by digital electronic circuits, computer hardware, firmware, software, and/or a combination thereof. Various embodiments of the inventive concept may be implemented as a data processing apparatus, for example, one or more programmable
20 processors and/or one or more computing apparatuses, or as a computer-readable recording medium and/or a computer program stored on the computer-readable recording medium. The computer program described above may be written in any programming language, including a compiled or interpreted language, and may be distributed in any form, such as a standalone program, module, subroutine, or the
25 like. The computer program may be distributed through a single computing apparatus, a plurality of computing apparatuses connected through the same network, and/or a plurality of computing apparatuses distributed to be connected through a plurality of different networks.

[00101] Meanwhile, embodiments disclosed in the specification may be
30 implemented in a form of a recording medium storing instructions executable by a computer. The instructions may be stored in a form of program codes, and, when executed by a processor, generate a program module to perform operations of the disclosed embodiments. The recording medium may be implemented as a computer-

readable recording medium. The computer-readable recording medium includes all kinds of recording media in which instructions capable of being decoded by a computer are stored. For example, there may be a ROM, a RAM, a magnetic tape, a magnetic disk, a flash memory, an optical data storage device, or the like.

5 **[00102]** The above description refers to detailed embodiments for implementing the inventive concept. The inventive concept may include embodiments in which a design is changed simply or which are easily changed, as well as the embodiments described above. In addition, the inventive concept may include technologies that are easily changed and implemented by using the above-
10 described embodiments. While the inventive concept has been described with reference to embodiments thereof, it will be apparent to those of ordinary skill in the art that various changes and modifications may be made thereto without departing from the spirit and scope of the inventive concept as set forth in the following claims

15 **[00103]** According to an embodiment of the inventive concept, compound-protein binding affinity may be accurately and effectively predicted based on compound-protein complex free.

[00104] Effects according to the inventive concept are not limited to the effects mentioned above, and other effects not mentioned will be clearly understood by those skilled in the art from the following description.

20 **[00105]** While the inventive concept has been described with reference to exemplary embodiments, it will be apparent to those skilled in the art that various changes and modifications may be made without departing from the spirit and scope of the inventive concept. Therefore, it should be understood that the above embodiments are not limiting, but illustrative.

25

WHAT IS CLAIMED IS:

1. A compound-protein binding affinity prediction method performed by at least one processor, the method comprising:

receiving compound data and protein data, which interact with each other;

5 generating an attribute vector of a compound and an attribute vector of a protein based on the input compound data and the input protein data;

calculating an attention value based on the attribute vector of the compound and the attribute vector of the protein;

generating a first interaction matrix based on the attention value;

10 learning a first AI model to predict a binding affinity and a non-covalent interaction of compound-protein by using the first interaction matrix as learning data; and

predicting the binding affinity and the non-covalent interaction of the compound-protein based on an output value of the first AI model.

15

2. The method of claim 1, wherein the receiving of the compound data includes:

generating a compound structure graph, and

wherein the compound structure graph has an attribute of an atom as a node

20 and has an attribute of a bond as an edge.

3. The method of claim 1, wherein the calculating of the attention value based on the attribute vector of the compound and the attribute vector of the protein includes:

25 calculating a key, a value, and a query based on the attribute vector of the compound and the attribute vector of the protein; and

crossing the calculated query of the compound and the calculated query of the protein, or crossing a key and a value of the compound and a key and a value of the protein to provide the crossed result to a sub-attention layer.

30

4. The method of claim 3, wherein the calculating of the attention value based on the attribute vector of the compound and the attribute vector of the protein further includes:

providing an output value of the sub-attention layer to a self-attention layer;
and
providing an output value of the self-attention layer to a feed-forward layer.

5 5. The method of claim 1, wherein the learning of the first AI model includes:

calculating an interaction score between one or more elements of the compound and one or more residues of the protein;

extracting a latent variable from the score; and

10 providing the latent variable to a fully connected layer.

6. The method of claim 1, wherein the output value of the first AI model includes the attribute vector of the compound, the attribute vector of the protein, the first interaction matrix, and the binding affinity predicted by the first AI model.

15

7. The method of claim 1, wherein the predicting of the binding affinity and the non-covalent interaction of the compound-protein based on the output value of the first AI model includes:

performing knowledge distillation by inputting the output value of the first
20 AI model into a second AI model; and

predicting the non-covalent interaction and the binding affinity in the second AI model, and

wherein data received by the first AI model is based on a compound-protein complex.

25

8. The method of claim 7, wherein the performing of the knowledge distillation includes:

calculating a loss function by comparing the first interaction matrix with a second interaction matrix generated by the second AI model.

30

9. A non-transitory computer readable recording medium including computer program to perform the compound-protein binding affinity prediction of claim 1.

10. A computing apparatus comprising:
- a communication module;
 - a memory; and
 - 5 at least one processor connected to the memory and configured to execute at least one computer-readable program included in the memory, wherein the at least one program includes instructions for:
 - receiving a structure of a compound and a structure of a protein, which interact with each other;
 - 10 generating an attribute vector of a compound and an attribute vector of a protein based on the input structure of the compound and the input structure of the protein;
 - calculating an attention value based on the attribute vector of the compound and the attribute vector of the protein;
 - 15 generating a first interaction matrix based on the attention value;
 - learning a first AI model to predict a binding affinity and a non-covalent interaction of compound-protein by using the first interaction matrix as learning data; and
 - predicting the binding affinity and the non-covalent interaction of the
 - 20 compound-protein based on an output value of the first AI model.

ABSTRACT

A compound-protein binding affinity prediction method performed by at least one processor includes receiving compound data and protein data, which interact with each other, generating an attribute vector of a compound and an attribute vector of a protein based on the input compound data and the input protein data, calculating an attention value based on the attribute vector of the compound and the attribute vector of the protein, generating a first interaction matrix based on the attention value, learning a first AI model to predict a binding affinity and a non-covalent interaction of compound-protein by using the first interaction matrix as learning data, and predicting the binding affinity and the non-covalent interaction of the compound-protein based on an output value of the first AI model.

FIG. 1

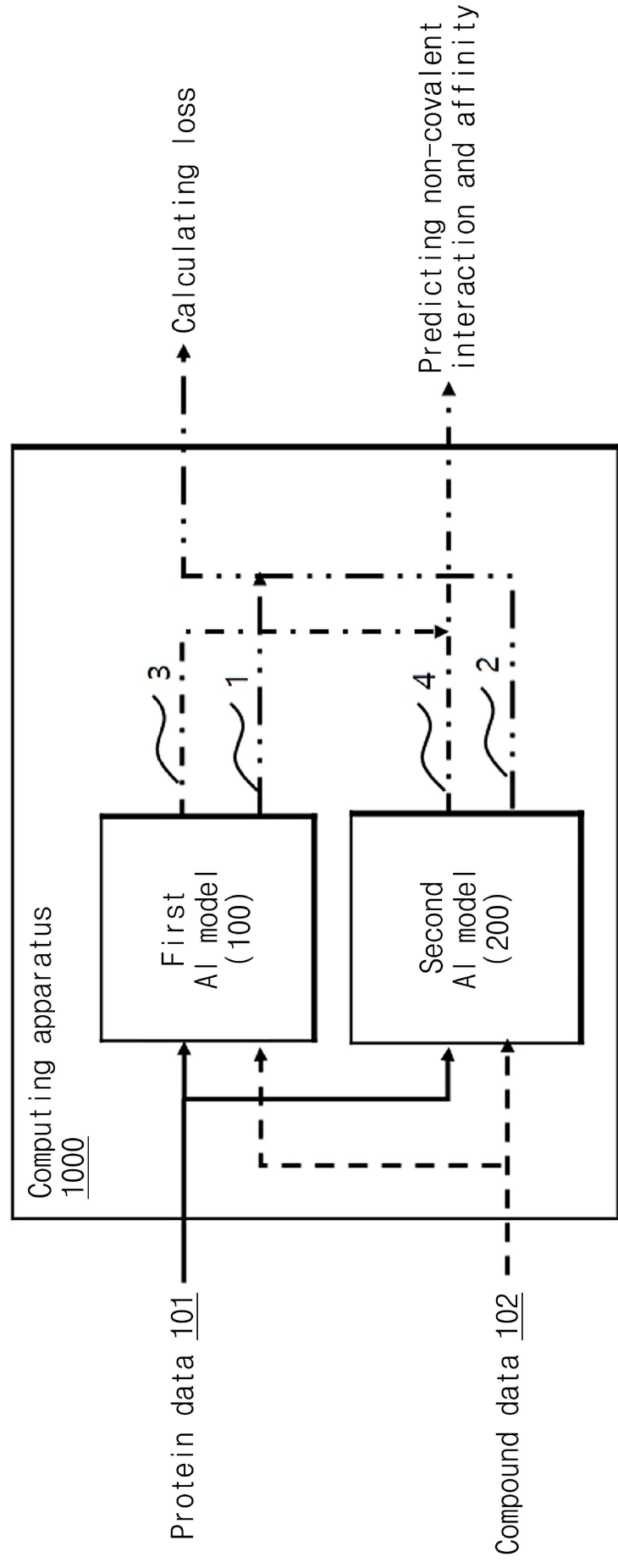


FIG. 2

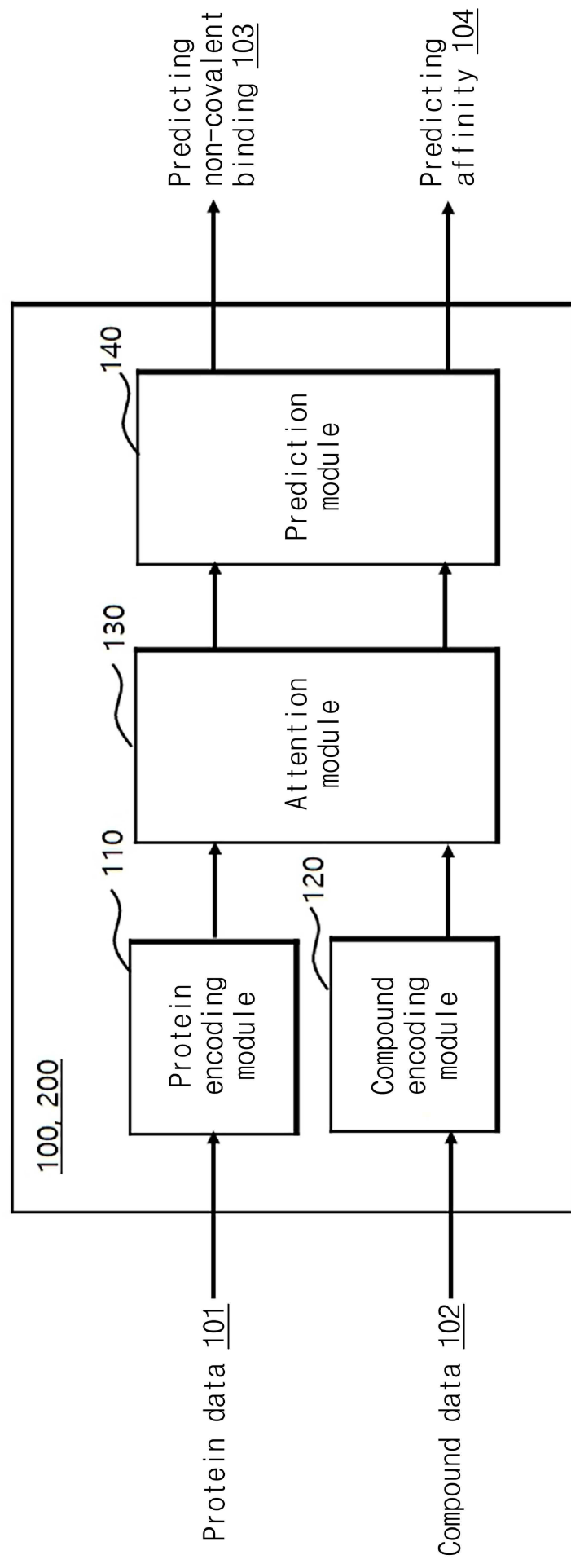


FIG. 3

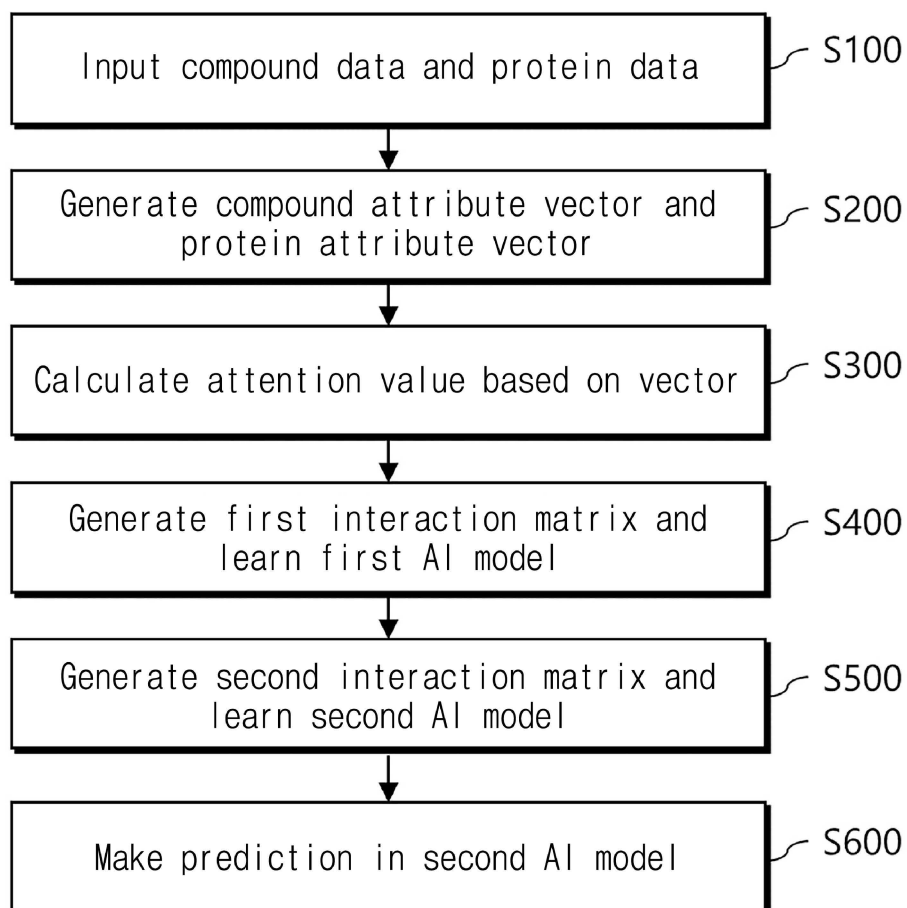


FIG. 4

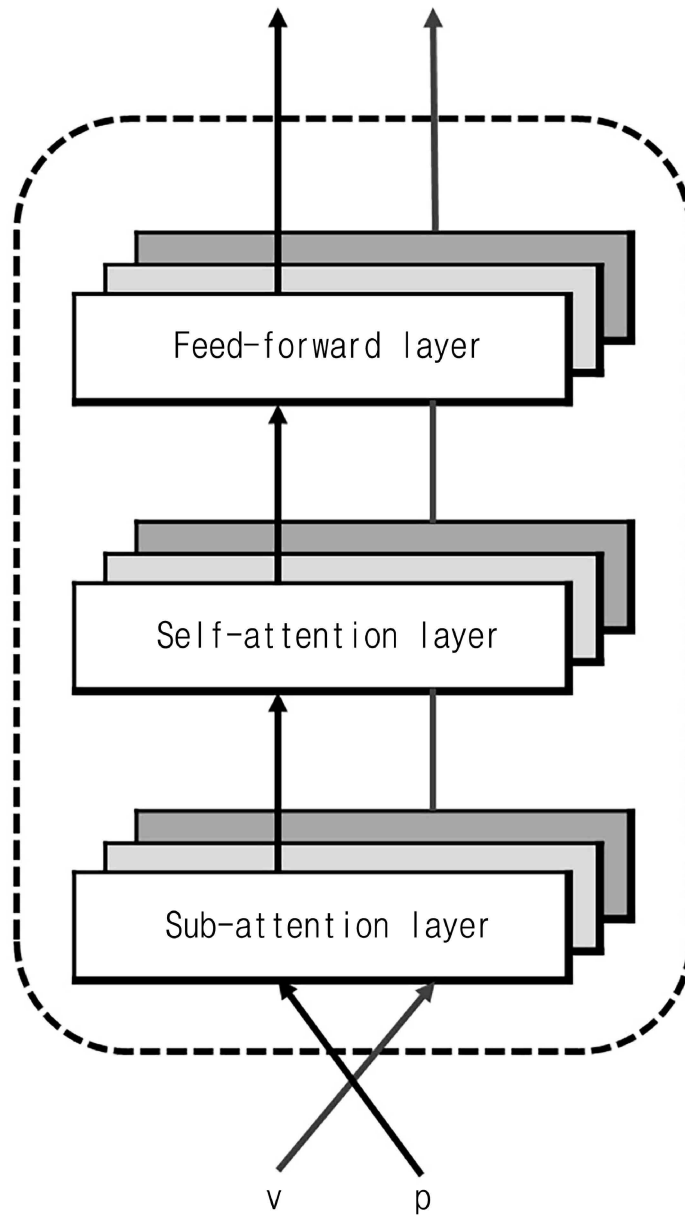


FIG. 5

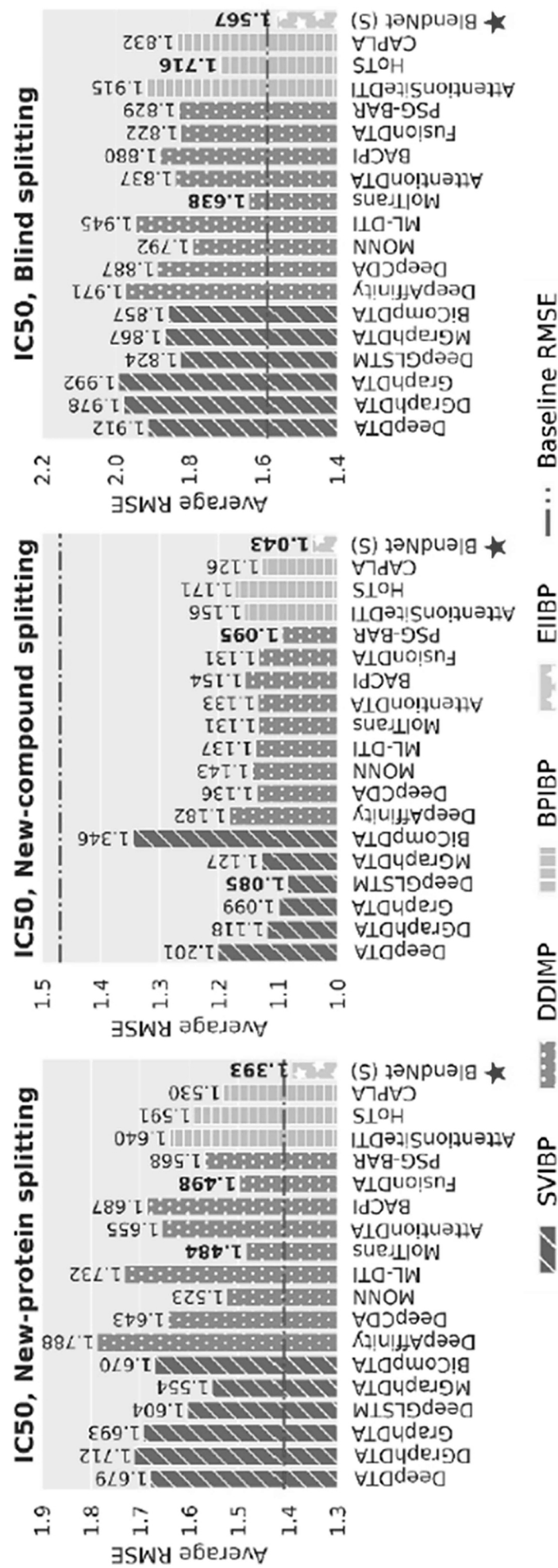
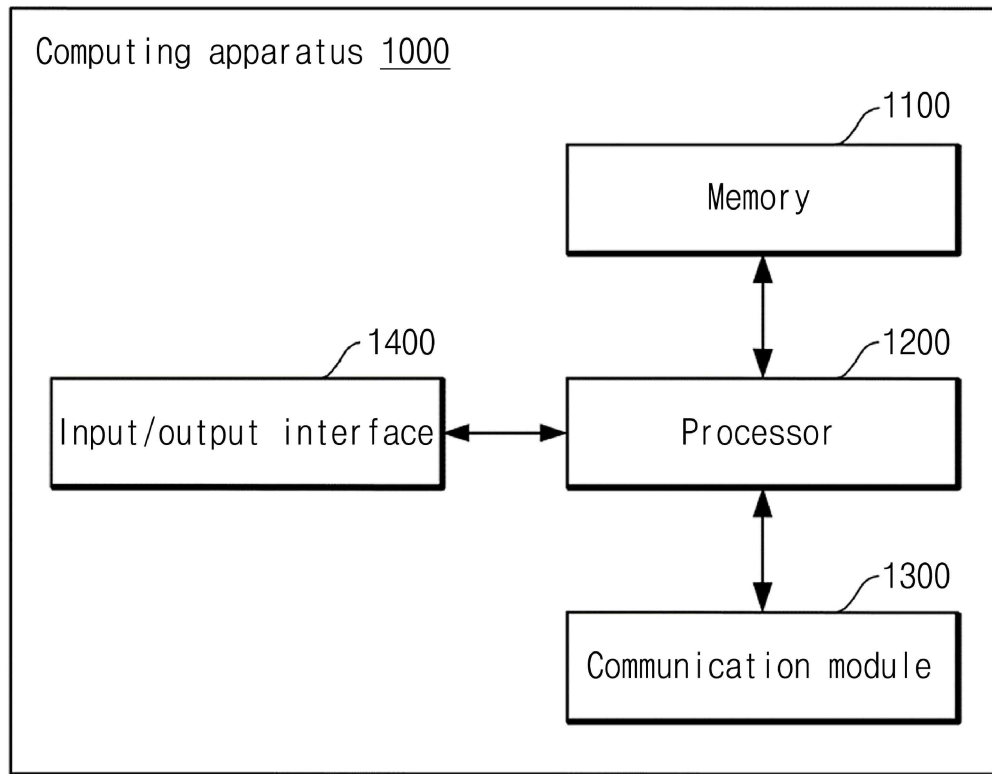


FIG. 6



Doc Code: ECOMM.AUTH/ECOMM.WTDW

Doc Description: Internet Communications Authorization/Internet Communications Authorization Withdrawal

PTO/SB/439 (11-15)

**AUTHORIZATION FOR INTERNET
COMMUNICATIONS IN A PATENT
APPLICATION OR REQUEST TO
WITHDRAW AUTHORIZATION FOR
INTERNET COMMUNICATIONS**

Application No.	
Filing Date	
First Named Inventor	
Art Unit	
Examiner Name	
Practitioner Docket No.	

To: Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

I. To authorize permission for Internet Communications.

☐ Recognizing that Internet communications are not secure, I hereby authorize the USPTO to communicate with the undersigned and practitioners in accordance with 37 CFR 1.33 and 37 CFR 1.34 concerning any subject matter of this application via video conferencing, instant messaging, or electronic mail. I understand that a copy of these communications will be made of record in the application file. (MPEP 502.03)

II. To withdraw authorization for Internet Communications.

☐ The authorization given on _____, to the USPTO to communicate with the undersigned and any practitioner in accordance with 37 CFR 1.33 and 37 CFR 1.34 concerning any subject matter of this application via Internet communications such as video conferencing, instant messaging, or electronic mail is hereby withdrawn. I understand that the withdrawal is effective when approved rather than when received.

I am the

☐ applicant.

☐ attorney or agent of record. Registration number _____.

☐ attorney or agent acting under 37 CFR 1.34. Registration number _____.

Signature

Date

Typed or printed name

Telephone Number

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Juristic entities must be represented by a patent practitioner (see 37 CFR 1.31, which is applicable to any paper filed on or after September 16, 2012, that is presented on behalf of a juristic entity, regardless of application filing date). Submit multiple forms if more than one signature is required, see below*.

☐ * Total of _____ forms are submitted.

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Request to Retrieve Electronic Priority Application(s)

Send completed form to: Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

COMPLETE IF KNOWN

Application Number

Filing Date

First Named Inventor

Art Unit

Examiner Name

Attorney Docket Number

Pursuant to 37 CFR 1.55(i), the undersigned hereby requests that the USPTO retrieve an electronic copy of each of the following foreign applications for which priority has been claimed under 35 U.S.C. 119(a)-(d) from a foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement. This Request must be submitted:

- within the later of sixteen months from the filing date of the prior foreign application or four months from the actual filing date of an application under 35 U.S.C. 111(a),
- within four months from the later of the date of commencement (37 CFR 1.491(a)) or the date of the initial submission under 35 U.S.C. 371 of an application entering the national stage under 35 U.S.C. 371, or
- with a petition under 37 CFR 1.55(e) or (f).

☐**OPTION A**

Please retrieve the priority application identified in Column C, a certified copy of which is contained in the EP or JP application identified in Columns A and B:

A	B			C	
Code for Participating Office (EP or JP only)	Application containing the non-participating priority application			Non-participating priority application to be retrieved	
	App. No.	Filing Date	Access Code	Country Code	App. No.
1					

☐**OPTION B**

This Request may be used for the infrequent circumstance when a claim for priority to an application filed in a participating foreign intellectual property office was made prior to that foreign intellectual property office becoming a participating foreign intellectual property office.

Please retrieve the priority application identified in Columns A and B:

A	B		
Code for Participating Office (e.g., EP) or WIPO DAS Depositing Office (e.g., AU, BR, CN, DK, EA, EE, ES, FI, GB, IB, IN, JP, KR, MA, NL, NZ, SE)	Application to be retrieved		
	App. No.	Filing Date	Access Code (for WIPO DAS Depositing Office)
1			
2			

The USPTO will not attempt to retrieve the identified priority application(s) unless an identical claim for foreign priority to the application identified above is made pursuant to 37 CFR 1.55(d) or a petition is granted under 37 CFR 1.55(e) or (f). Applicants are advised to consult Private PAIR (accessed through www.uspto.gov) to assure that the retrieval has been successful. The applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period set forth in 37 CFR 1.55(g)(1).

I hereby declare that I have the authority to grant access to the above-identified foreign application(s).

Signature

Date

Printed or Typed Name

Telephone Number

Title

Registration Number, if applicable

This collection of information is required by 37 CFR 1.55(d). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)	* 4422. 00. , 22 / 04		
	, 00 - \$. 0		
	020, . 2 00 * . 00. 004	- * , . * 24 , 4 * - .	
	* 404 . 0		
	3. 2 0 04 , . 2 0		
	* 004 00 \$ 02/ 00 , 22 / 04	444\$23/ \$0	

U.S.PATENTS

3. 2 0 04 - 00 * . 00 20 , 0	4. 00. 0 , 22 / 04	. 0 0 - 0000	*2220 \$. 0	. 2 0 004. 00. 00 04* 4422. . 0 002000 \$ 0222 0. 0	4. - 024 022 . 24* 0 02 / 1040 - 020. . 04. 22. - 02 04- 020. . 0 , 0 2402 * 440. 4
0					

*0002 / 021 00. 00 . 000000 . . 24 0 04 . 00. 020 00. 0 042 . 00. 42. 20 222/ 010 * 00 / 2000. 0

U.S.PATENT APPLICATION PUBLICATIONS

3. 2 0 04 - 00 * . 00 20 , 0	42/ 22. 00. , 22 / 04	. 0 0 - 0000	42/ 22. 00. \$. 0	. 2 0 004. 00. 00 04* 4422. . 0 002000 \$ 0222 0. 0	4. - 024 022 . 24* 0 02 / 1040 - 020. . 04. 22. - 02 04- 020. . 0 , 0 2402 * 440. 4
0					

*0002 / 021 00. 00 . 000000 . . 24 0 04 2/ 22100 * 4422. 00. 20 00. 0 042 . 00. 42. 20 222/ 010 * 00 / 2000. 0

FOREIGN PATENT DOCUMENTS

3. 2 0 04 - 00 * . 00 20 , 0	, 0400. \$ 0222 0. 0 , 22 / 04	- 02. 00 - 000/ 0	. 0 0 - 0000	42/ 22. 00. \$. 0	. 2 0 004. 00. 00 04 * 4422. . 0002000 \$ 0222 0. 0	4. - 024 022 . 24* 0 02 / 1040 - 020. . 0 4. 22. - 02 04- 020. . 0 , 0 2402 * 440. 4	- 0
0	022/ 2/ 0222/ 0020	. -	*	/ 2/ 022/ 2 4	- 0022, . 00. . 24 . 004200 - 0 \$- , 02. 0. 00. 00. 2		<input checked="" type="checkbox"/>
/	022/ 2/ 0222\$/ 340	. -	*	/ 2/ 0224200	\$, * - . . , , * , - 0		<input checked="" type="checkbox"/>

*0002 / 021 00. 00 . 000000 . . 2, 0400. 4. 00. 0\$ 0222 0. 020 00. 0 042 . 00. 42. 20 222/ 010 * 00 / 2000.

NON-PATENT LITERATURE DOCUMENTS

Not for submission under 37 CFR 1.99)

Not for submission under 37 CFR 1.99)

A° °™1t1z, u1 z8°		
, °™1, 1t8		
, °š, 118, 1v8zt1°	cA, ÜV1, , A, 1	
A°t1 z t		
èð11 z8°, 118		
Att1°z81, 11×8t, u1 z8°		, , , , mÿz, ,

<p>èð11 z8° z t 1 5</p>	<p>4 t8 , 1</p>	<p>z1w, 8 z11 8 1, t, 8 1ut, 1° pz 4 A, 1 A, èí í è, ct¹tt8 1, t, 8 1°t18°pw, 8z 1°°1°°1t8t¹tt8 1, t, 8 t81 pz 11x¹¹ 111z z8¹¹u°z¹¹S8¹¹S11¹¹Šu1¹¹1t1¹¹¹8t1¹¹, 1t8¹¹118Sp¹v1¹¹1 8úŠŠu8 zu1 z8°p¹¹ °uz¹¹S, 8¹¹1t1 1z, /1¹¹11uzt¹¹w, 8°8°uz¹¹S, 8, .</p>	<p>í ¥</p>
	<p>1</p>	<p>c1z11 z c 81 8t 1™"˘ z, z1 1,, z t1°°8, 1t1z, 1°°°1t8 z-##8/ &" ' ##! &# % %% '%%"#&" " %%% #&" %%&% &" # '%& &#%! #&# '%#%'& &# ! "-"-!8!</p>	<input type="checkbox"/>
		<p>%% % # %&% %&#, ## # #&# %'%%' &'%"#&#& &" ' &!' -' #&%'& &" ' #& &# '! #/% #8/" & %& # ###! #& #8/ # % ! &# % #&# # %&# ! 3!</p>	<input type="checkbox"/>

%&! \$#&'&% % ##&"%#& "&-'%%' #'&#!#% &! "%' #'&#& #'&#& #'&#& ' #'&#& ##'%'#& %&'&#&

EXAMINER SIGNATURE

##% # %#%#% %! #%	# %%#&" &# %&%
-------------------	----------------

*MCAelč MR: Initial if reference consiÁereÁÁwÁetÁer or not citation is in conformance witÁ œPMP 6Á9. I raw line tÁrougÁ a citation if not in conformance anÁ not consiÁereÁ. IncluÁe copy of tÁis form witÁ ne## communication to applicant.

1 See KinA CoAes of ~ SPMI Patent I ocuments at www.~ SPMI.~ d v or cePMP 9A1.4A. 2 Mnter office tAat issueA tAe A documentA by tAe two-letter coAe (WIPd StanAarA SM3). 3 Or Japanese patent A documentsA tAe inAication of tAe year of tAe reign of tAe Mmperor must precede tAe serial number of tAe patent A document. 4 KinA of A document by tAe appropriate symbols as inAicateA on tAe A document unAer WIPd StanAarA SM16 if possible. 5 Applicant is to place a cAeck mark Aere if MnglisA language translation is attacAeA

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)

Application č umber		
Öling I ate		
Örst č ameÁ Inventor	SAč - í Y~ č PARK	
Art ~ nit		
M#aminer č ame		
Attorney I ocket č umber	PP68Á7284	

CERTIFICATION STATEMENT

Please see 37 CÖR 1.97 anÁ 1.98 to make tÁe appropriate selection(s):

☐ MÁat eacÁ item of information containeÁ in tÁe information Áisclosure statement was first citeÁ in any communication from a foreign patent office in a counterpart foreign application not more tÁan tÁree montÁs prior to tÁe filing of tÁe information Áisclosure statement. See 37 CÖR 1.97(e)(1).

OR

☐ MÁat no item of information containeÁ in tÁe information Áisclosure statement was citeÁ in a communication from a foreign patent office in a counterpart foreign applicationÁanÁÁto tÁe knowleÁge of tÁe person signing tÁe certification after making reasonable inquiryÁno item of information containeÁ in tÁe information Áisclosure statement was known to any inÁiviÁual ÁesignateÁ in 37 CÖR 1.56(c) more tÁan tÁree montÁs prior to tÁe filing of tÁe information Áisclosure statement. See 37 CÖR 1.97(e)(2).

☐ See attacÁeÁ certification statement.

☐ MÁe fee set fortÁ in 37 CÖR 1.17 (p) Áas been submitteÁ ÁerewitÁ.

☒ A certification statement is not submitteÁ ÁerewitÁ.

SIGNATURE

A signature of tÁe applicant or representative is requireÁ in accorÁance witÁ CÖR 1.33Á1Á18. Please see CÖR 1.4(Á) for tÁe form of tÁe signature.

Signature	/Byungwoong Park/	I ate (YYYY-œœ-I)	2Á25-12-Á6
č ame/Print	Byungwoong Park	Registration č umber	8ÁÁ134

MÁis collection of information is requireÁ by 37 CÖR 1.97 anÁ 1.98. MÁe information is requireÁ to obtain or retain a benefit by tÁe public wÁicÁ is to file (anÁ by tÁe ~ SPMđ to process) an application. ConfiÁentiality is governeÁ by 35 ~ .S.C. 122 anÁ 37 CÖR 1.14. MÁis collection is estimateÁ to take 1 Áour to completeÁincluÁing gatÁeringÁpreparing anÁ submitting tÁe completeÁ application form to tÁe ~ SPMđ. Mme will vary ÁepenÁing upon tÁe inÁiviÁual case. Any comments on tÁe amount of time you require to complete tÁis form anÁ/or suggestions for reÁucing tÁis burÁenÁsÁoulÁ be sent to tÁe CÁief Information đ fficerÁ ~ .S. Patent anÁ MraÁemark đ fficeÁ ~ .S. I epartment of CommerceÁP.đ. Bo# 145ÁÁAle#anÁriaÁVA 22313-145Á. I đ č đ MSMđ I ÖMMS đ R Čđ œPhMMI Čđ RœS Mđ MI IS AI I RMSS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.h. 93-579) requires that you be given certain information in connection with your submission of the attached form relating to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission relating to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.
2. A record from this system of records may be disclosed as a routine use in the course of presenting evidence to a court, magistrate or administrative tribunal including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed as a routine use to a member of Congress submitting a request involving an individual to whom the record pertains when the individual has requested assistance from the member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed as a routine use to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974 as amended pursuant to 5 U.S.C. 552a(m).
5. A record relating to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed as a routine use to the International Bureau of the World Intellectual Property Organization pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed as a routine use to another federal agency for purposes of national security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed as a routine use to the Administrator, General Services Administration or designee during an inspection of records conducted by - SA as part of that agency's responsibility to recommend improvements in records management practices and programs under authority of 44 U.S.C. 29A4 and 29A6. Such disclosure shall be made in accordance with the - SA regulations governing inspection of records for this purpose and any other relevant (i.e., - SA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed as a routine use to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Either the record may be disclosed subject to the limitations of 37 CFR 1.14 as a routine use to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and a new application is referenced by either a published application or an application open to public inspections or an issued patent.
9. A record from this system of records may be disclosed as a routine use to a Federal, State or local law enforcement agency if the SPM becomes aware of a violation or potential violation of law or regulation.

**COMBINED DECLARATION (37 CFR §1.63) FOR UTILITY PATENT
APPLICATION AND ASSIGNMENT FORM¹**

The undersigned acknowledges that this document is being used both as an assignment of the invention and as the declaration (37 CFR 1.63) for a Utility or Design Application.

WHEREAS, the undersigned, hereinafter referred to collectively as Assignor, has invented:

Title: METHOD FOR PREDICTING COMPOUND-PROTEIN BINDING AFFINITY
AND APPARATUS THEREOF

for which Assignor is about to make or has made United States or International application for patent.

WHEREAS, as a below named inventor(s), I/(we) hereby declare that:

Section I. Declaration

This declaration is directed to:

- ☒ The attached U.S. non-provisional patent application, or
- ☐ U.S. non-provisional patent application number _____, filed on _____, or
- ☐ PCT international patent application number _____, filed on _____.

The above-identified application was made or authorized to be made by me (us).

I (we) believe that I am (we are) the original inventor (original joint inventors) of a claimed invention in the above-identified application.

I (we) hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

I hereby state that I have reviewed and understood the contents of the above identified application, including the claims.

¹ This form requires the use of an Application Data Sheet.

I acknowledge that I am aware of the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56.

Section II. Assignment

In consideration of the sum of One Dollar (\$1.00) and other good and valuable consideration paid to each of the undersigned, the undersigned hereby sell(s) and assign(s) to

UIF (University Industry Foundation), Yonsei University

having an address at 50 Yonsei-ro, Seodaemun-gu, Seoul, 03722, Republic of Korea (hereinafter designated as the Assignee), the entire (100%) right, title and interest for the United States as defined in 35 USC §100, in the invention described in the application identified in Section I of this document.

Assignor hereby confirms any prior assignment to Assignee, and to the extent that Assignor has not already done so, agrees to assign, and hereby does, sell, assign and transfer unto Assignee and its successors in interest, the full and exclusive right, title and interest in the United States of America and throughout the world, including the right to claim priority under the laws of the United States, the Paris Convention, and any foreign countries, to the inventions as described in the aforesaid application, to the aforesaid application itself, and all divisions, continuations, continuations-in-part, or other applications claiming priority directly or indirectly from the aforesaid application, and any United States or foreign Letters Patent, utility model, or other similar rights which may be granted thereon, including reissues, reexaminations and extensions thereof, and all copyright rights throughout the world in the aforesaid application and the subject matter disclosed therein, these rights, title and interest to be held and enjoyed by Assignee to the full end of the term for which the Letters Patent, utility model, or other similar rights, are granted and any extensions thereof as fully and entirely as the same would have been held by Assignor had this assignment and sale not been made, and the right to sue for, and recover for past infringements of, or liabilities for, any of the rights relating to any of the applications, patents, utility models, or other similar rights, resulting therefrom, and the copyright rights;


Assignor hereby covenants and agrees to execute all instruments or documents required or requested for the making and prosecution of any applications of any type for patent, utility model, or other similar rights, and for copyright, in the United States and in all foreign countries including, but not limited to, any provisional, continuation, continuation-in-part, divisional, renewal or substitute thereof, and as to letters patent any reissue, re-examination, or extension thereof, and for litigation regarding, or for the purpose of protecting title and to the said invention, the United States application for patent, or Letters Patent therefor, and to testify in support thereof, for the benefit of Assignee without further or other compensation than that above set forth;


Assignor hereby covenants that no assignment, sale, license, agreement or encumbrance has been or will be entered into which would conflict with this Assignment; and

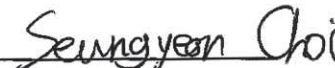
Assignor hereby requests the United States Patent and Trademark Office to issue the Letters Patent of the United States of America to Assignee, and requests that any official of any country or countries foreign to the United States, whose duty it is to issue or grant patents and applications as aforesaid, to issue the Letters Patent, Utility Model Registration or other similar right to Assignee.

The undersigned hereby grant(s) the law firm of **ZION IP Law PLLC** the power to insert on this Declaration and Assignment any further identification which may be necessary or desirable in order to comply with the rules of the U.S. Patent and Trademark Office for recordation of this document.

Section III. Inventor(s)/Assignor(s) Signature(s)

LEGAL NAME OF INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	PARK	SANGHYUN	
SIGNATURE <u></u> DATE <u>25.12.3</u>			

LEGAL NAME OF INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	SEO	SANGMIN	
SIGNATURE <u></u> DATE <u>25/12.3</u>			

LEGAL NAME OF INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	CHOI	SEUNGYEON	
SIGNATURE <u></u> DATE <u>25.12.3</u>			

LEGAL NAME OF INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	KIM	HWANHEE	

SIGNATURE Kim Hwanhee DATE 25.12.03

LEGAL NAME OF INVENTOR/ASSIGNOR	Family Name	First Given Name	Second Given Name
	LEE	JIEUN	

SIGNATURE Lee Jieun DATE 25.12.03