

Communication

AI-Powered Western Blot Interpretation: A Novel Approach to Studying the Frameshift Mutant of Ubiquitin B (UBB⁺¹) in Schizophrenia

Artur Fabijan ^{1,*}, Michał Chojnacki ^{2,†}, Agnieszka Zawadzka-Fabijan ³, Robert Fabijan ⁴, Michał Piątek ⁵, Krzysztof Zakrzewski ¹, Emilia Nowosławska ¹ and Bartosz Polis ¹

- ¹ Department of Neurosurgery, Polish-Mother's Memorial Hospital Research Institute, 93-338 Lodz, Poland; krzysztof.zakrzewski@iczmp.edu.pl (K.Z.); emilia.nowoslawska@iczmp.edu.pl (E.N.); jezza@post.pl (B.P.)
- ² Department of Medical Biochemistry, Medical University of Lodz, 92-215 Lodz, Poland; chojnacki.michal84@gmail.com
- ³ Department of Rehabilitation Medicine, Faculty of Health Sciences, Medical University of Lodz, 90-419 Lodz, Poland; agnieszka.zawadzka@umed.lodz.pl
- ⁴ Independent Researcher, Luton LU2 0GS, UK; robert.f.fabijan@gmail.com
- ⁵ Babinski Memorial Hospital, 91-229 Lodz, Poland; piatekmichal1990@gmail.com
- * Correspondence: artur8944@wp.pl
- † These authors contributed equally to this work.

Abstract: The application of artificial intelligence (AI) in the analysis of molecular biology data is becoming increasingly widespread. The Western Blot (WB) technique, a cornerstone in proteomic research, facilitates the identification and analysis of proteins, such as the frameshift mutant of ubiquitin B (UBB⁺¹). In our study, we attempted to assess the potential of four different AI models—Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4—in the analysis of WB imagery containing UBB⁺¹, derived from peripheral blood studies of patients suffering from schizophrenia. Participants, all male and diagnosed with schizophrenia, were recruited from the Specialist Psychiatric Care Team of Babinski Hospital in Lodz. After obtaining their informed consent, blood samples were collected and transported to the laboratory of the Department of Medical Biochemistry at the Medical University of Lodz. The samples were processed, synthesis of Ub-⁴⁸UBB⁺¹ dimers was performed, and the WB technique was applied. The result of the WB analysis, in the form of a photograph with basic labels but without a legend (JPG format), was implemented into ChatGPT 4, Microsoft Copilot, Gemini and Gemini Advanced. Following the implementation of the image, the command ‘Could you analyze the attached photo?’ was added, along with the protocol from Sample Preparation and Synthesis of Ub-⁴⁸UBB⁺¹ Dimers. The AI models effectively analyzed and interpreted WB images, with variations in their approaches and depth. Gemini excelled in detailing the WB process and biological significance of bands, while Gemini Advanced focused on specific band identification, especially Ub-⁴⁸UBB⁺¹ dimers. Microsoft Copilot provided a basic overview with less technicality, and ChatGPT 4 offered comprehensive band interpretations, linking them to patient samples and standards, thus confirming the hypothesis about the differing capabilities of these models. This discovery demonstrates the advanced capabilities of ChatGPT 4 and highlights the growing role of AI in scientific research, including the interpretation of results.

Keywords: Microsoft Copilot; Gemini; ChatGPT 4; Western Blot; ubiquitin; UBB⁺¹; artificial intelligence; schizophrenia



Citation: Fabijan, A.; Chojnacki, M.; Zawadzka-Fabijan, A.; Fabijan, R.; Piątek, M.; Zakrzewski, K.; Nowosławska, E.; Polis, B. AI-Powered Western Blot Interpretation: A Novel Approach to Studying the Frameshift Mutant of Ubiquitin B (UBB⁺¹) in Schizophrenia. *Appl. Sci.* **2024**, *14*, 4149. <https://doi.org/10.3390/app14104149>

Academic Editor: Zhonghua Sun

Received: 8 April 2024
Revised: 10 May 2024
Accepted: 11 May 2024
Published: 14 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The application of artificial intelligence (AI) in the analysis of molecular biology data is becoming increasingly widespread, as reflected in numerous studies and practical applications [1,2]. A significant example of AI application in molecular biology is the analysis of

results obtained using the Western Blot (WB) technique, one of the fundamental research methods in proteomics [3]. Despite its utility, WB analysis faces several challenges, including variability in gel electrophoresis conditions, differences in antibody specificity and affinity, and the subjective interpretation of results. These factors can lead to inconsistencies in data analysis, underscoring the need for more objective and reproducible approaches.

In recent years, there has been a heightened interest in utilizing AI models in various fields, including medicine [4]. The advancement of AI technology, especially following the release of the ChatGPT model version 4 by OpenAI, has significantly contributed to progress in research on the application of language models in the biological sciences sector [5]. The recent enhancement of ChatGPT's capabilities to include image analysis [6] has opened new perspectives in the interpretation of biological data. However, there is still a need for further research into the use of these models in the analysis of molecular research results.

Microsoft Copilot, unveiled by Microsoft on 7 February 2023, represents an innovative implementation of an advanced language model in the form of a chatbot. This system can generate creative content, including poetry and music, utilizing the Suno AI plugin [7]. Its introduction marked a pivotal element in Microsoft's strategy for artificial intelligence, replacing its earlier digital assistant, Cortana [8]. From 21 September 2023, as part of a rebranding process, all versions of this tool were unified under the name Microsoft Copilot. In December 2023, Copilot was integrated into numerous Windows 11 installations and a limited number of Windows 10 systems. The expansion of Copilot's functionalities emphasizes the growing role of artificial intelligence in Microsoft's product offerings and its significance in future technological development.

The Gemini and Gemini Advanced models, developed by Google AI, constitute two distinct iterations in the realm of large language models (LLMs), differing in terms of size, capacity, and scope of capabilities. With 137 billion parameters, the Gemini model is a smaller model designed for rapid text generation and efficient operation on devices with limited computational power. Its primary applications include generating responses to questions, creating text summaries, language translation, and crafting brief creative content. This version focuses on the fundamental functions of LLMs, offering greater accessibility and operational speed [9].

Meanwhile, Gemini Advanced is a significantly larger model, providing higher precision and detail in text generation. This model extends the capabilities of standard LLM tasks, including writing code, generating scripts, creating music, and composing emails and letters. Gemini Advanced targets more complex applications requiring deep contextual understanding and advanced content analysis [10].

The WB technique, pivotal in proteomic research, allows for the detection and analysis of proteins, including the mutated ubiquitin protein UBB^{+1} , which arises from an incorrect reading of the genetic code, leading to the formation of a specific molecular "tail" [11]. The significance of this technique is particularly important in the context of schizophrenia, as numerous studies indicate a connection between ubiquitin dysfunctions and the pathomechanisms of this disease [12]. Addressing the challenges in WB analysis, such as variability and subjectivity, with AI could significantly enhance the accuracy and reliability of these studies.

In our study, we attempted to assess the potential of four different AI models—Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4—in the analysis of WB imagery containing UBB^{+1} , derived from peripheral blood studies of patients suffering from schizophrenia. The central focus of the research was the hypothesis that all these AI models would be capable of conducting fundamental analysis and interpretation of the WB image, identifying key elements such as the presence and characteristics of individual protein bands, and understanding the fundamental steps of the WB procedure upon receiving a detailed description of its execution. However, we hypothesized that the degree of accuracy, depth of technical details, and the biological context in applying WB results may vary among the models. This study aimed to evaluate to what extent advanced AI tools can support

analysis and interpretation of data in biomedical research, particularly in the context of complex techniques like WB. Incorporating AI into the interpretation of molecular biology data presents a promising avenue towards more objective analysis. By leveraging AI technologies, we can significantly reduce the dependency on subjective human observation, potentially minimizing biases inherent in manual data interpretation. AI algorithms are adept at consistently applying predefined criteria to analyze and interpret complex datasets obtained from techniques such as WB. As AI technology continues to evolve, its application in the field of molecular biology promises to revolutionize our approach to data interpretation, opening up new possibilities for discoveries and advancements in the field.

2. Materials and Methods

The study participants were recruited within the Specialist Psychiatric Care Team of the Babinski Hospital in Lodz from June 2018 to December 2019. The entire procedure was approved by the Bioethics Committee of the Medical University of Lodz in accordance with decision number RNN/208/17/KE. Each participant expressed informed consent to participate by signing the appropriate form and agreed to use their biological material for molecular research purposes.

The study exclusively included men previously diagnosed with schizophrenia, aged over 18 years, who voluntarily agreed to participate ($n = 32$). The control group consisted of men who were not diagnosed with a chronic disease, also over 18, who expressed their willingness to participate in the study ($n = 8$). Exclusion criteria included: age below 18 years, a primary psychiatric diagnosis other than schizophrenia, severe neurological and medical pathologies, an unstable clinical condition of the patient, severe symptoms of Parkinson's disease, significant sensory impairments, acute cognitive disorders, lack of signed informed consent to participate in the study, or lack of understanding of the information regarding the objectives and conditions of the study. After signing informed consent for participation in the study and consent to use biological material for molecular research, peripheral venous blood samples on EDTA were collected and transported to the laboratory of the Department of Medical Biochemistry, Medical University of Lodz.

2.1. Sample Preparation

Venous blood samples from patients were collected in tubes containing K3-EDTA. The collected samples (600 μ L) were homogenized using a TissueRuptor homogenizer (QIAGEN, Hilden, Germany) at maximum speed for 30 s. Subsequently, the samples were centrifuged at $18,000 \times g$ at 4 °C for 30 min. The supernatant was collected, and the samples were concentrated with the addition of trichloroacetic acid (20%). After acid addition, the samples were incubated at 4 °C for about 30 min and then centrifuged at $13,000 \times g$ for 15 min at 4 °C. The supernatant was removed and the protein pellet was resuspended in 500 μ L of acetone. After acetone washing, the samples were centrifuged at $13,000 \times g$ for 15 min at 4 °C. The supernatant was removed, and the pellet was resuspended in 75 μ L of $2 \times$ PLD (protein loading dye).

A 10 μ L amount of each sample was subjected to electrophoresis in a 15% SDS-PAGE gel. After electrophoretic separation, the proteins were transferred to a nitrocellulose membrane (Thermo Scientific, Waltham, MA, USA) in a Mini-PROTEAN chamber (Bio-Rad, Tokyo, Japan) for 60 min at a constant voltage of 100 V. The prepared membranes were washed for 3×5 min in TBST buffer. After washing the membranes, the proteins were incubated with the primary anti-Ub⁺¹ antibody (40B3) (Santa-Cruz, Santa Cruz, CA, USA) at a dilution of 1:1000 for 60 min at room temperature with gentle shaking. After the incubation, the membranes were washed for 3×5 min in TBST buffer and incubated with the secondary mouse antibody Peroxidase AffiniPure Goat IgG (H + L) (Biokom, Pécs, Hungary) at a dilution of 1:50,000 for 45 min with gentle shaking. After the incubation, the membrane was washed for 3×5 min in TBST and then placed in ECL Western Blotting Substrate (Pierce™, Waltham, MA, USA) solution to induce a luminescence reaction and

protein detection. After incubation (5 min), detection and visualization were carried out using a Chemi-Doc device (Bio-Rad, Tokyo, Japan).

2.2. Synthesis of Ub-⁴⁸UBB⁺¹ Dimers

For the synthesis of Ub-⁴⁸UBB⁺¹ dimers, the following proteins were used: His-UBB⁺¹ and Ub^{K48R/K63R}. The enzymatic synthesis was carried out using enzymes 500 nM Uba1 and 20 μM E2-25K (Boston Biochem, Cambridge, MA, USA), in reaction buffer (50 mM TRIS pH 8.0, 15 mM MgCl₂, 20 mM creatine phosphate 1.2 U/mL inorganic yeast pyrophosphatase, 1.2 U/mL creatine phosphokinase), 20 mM ATP with the addition of 4 mM TCEP according to the protocol [13]. The reaction was carried out at 30 °C for 24 h. After the enzymatic reaction, the obtained dimers were purified on a HisTrap 5 mL column (Cytiva, Tokyo, Japan) in affinity chromatography binding buffer (20 mM PB, 200 mM NaCl, 10 mM imidazole, pH 7.4). In affinity chromatography elution buffer (20 mM PB, 200 mM NaCl, 250 mM imidazole, pH 7.4) gradient, the Ub-⁴⁸UBB⁺¹ fraction was obtained. To refine the purification process, the dimer was subjected to chromatographic purification using a Superdex 75 16/60 column (Cytiva, Tokyo, Japan) in PBS pH 7.4. The presence of the reaction product was confirmed by SDS-PAGE polyacrylamide gel electrophoresis.

2.3. Application of the AI Model

The result of the WB analysis, in the form of a photograph with basic labels but without a legend (JPG format), was implemented into ChatGPT 4, Microsoft Copilot, Gemini and Gemini Advanced. Following the implementation of the image, the command ‘*Could you analyze the attached photo? As additional information, I am attaching the examination protocol.*’ was added, along with the protocol from Sample Preparation and Synthesis of Ub-⁴⁸UBB⁺¹ Dimers. The entire analysis is available in File S1 in Supplementary Materials. The selection of these AI models was strategic, aiming to leverage their distinct capabilities and areas of expertise. ChatGPT 4 and Microsoft Copilot were chosen for their proven track record and extensive documentation in scientific literature, which has established them as reliable tools for complex data interpretation [6,14,15]. On the other hand, the inclusion of the Gemini and Gemini Advanced models was motivated by their novelty and the claims of their developers regarding their advanced analytical capabilities [16]. By evaluating both established and emerging AI models, our study aimed to provide a comprehensive assessment of their utility in enhancing the interpretation of molecular biology data, particularly in the nuanced analysis required for WB imagery.

2.4. Selection Criteria for AI Tools

The selection of AI models for this study, specifically Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4, was primarily driven by their distinct capabilities and the availability of recent enhancements in these models. Our decision to include Gemini and Gemini Advanced was influenced by their novelty and the recent integration of capabilities that allow these models to process and interpret image data—a crucial requirement for the analysis of WB imagery. Gemini, being a new model released this year, presents an opportunity to assess the latest advancements in AI technologies and their application to complex biomedical data.

Our criteria for selecting these models also considered the practical need for AI tools that can interpret visual and textual data effectively, given the complex nature of WB analyses which often involves the interpretation of both image patterns and contextual information. The inclusion of Microsoft Copilot and ChatGPT 4 was based on their proven proficiency in handling and interpreting extensive datasets and generating detailed, contextually accurate textual outputs, respectively.

3. Results

3.1. Results of the WB Analysis

Figure 1 shows the result of a WB analysis focused on ubiquitin conjugates (Ub) and UBB⁺¹ protein (Ub-⁴⁸UBB⁺¹), prepared from cellular lysates. The samples underwent a process involving lysis, centrifugation, TCA protein precipitation, denaturation, SDS-PAGE electrophoresis, transfer to a nitrocellulose membrane, TBST blocking, incubation with primary and secondary antibodies, and chemiluminescent detection.

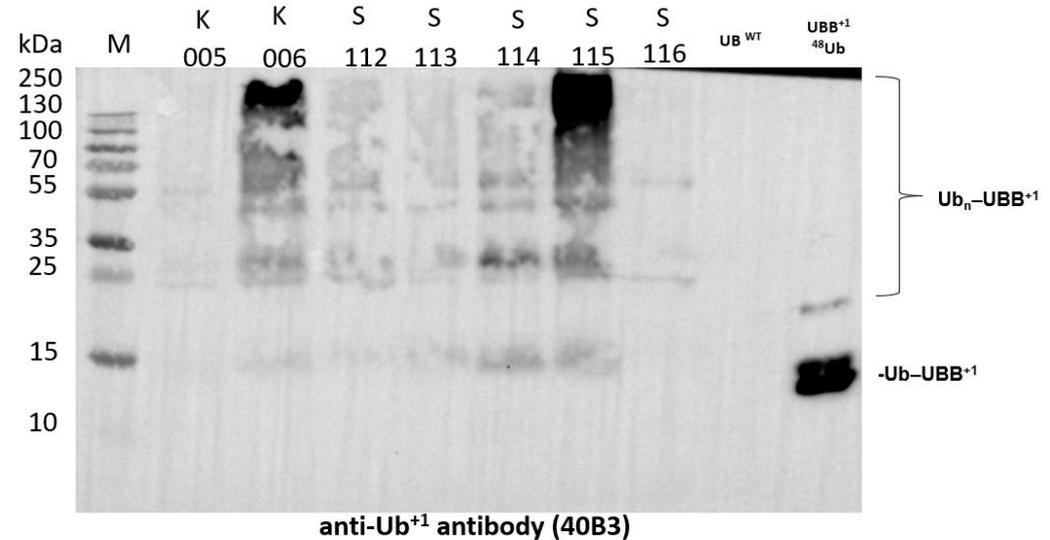


Figure 1. WB analysis results: peripheral blood proteins in schizophrenia patients and control group. This analysis compares the control group (K), numbered 005–006, with the schizophrenia patient group (S), numbered 112–116. The molecular weight marker is labeled ‘M’. Notably, Ub-⁴⁸UBB⁺¹ dimer bands are evident on the right, with an absence in the Ub^{WT} (wild type) section. The synthesized Ub-⁴⁸UBB⁺¹ dimer was also detected, as a positive control for antibodies. The use of anti-Ub⁺¹ (40B3) antibodies revealed a significant presence of polymeric UBB⁺¹-Ub in both schizophrenia (S) and control (K) samples. The labels ‘UBB⁺¹’ and ‘⁴⁸Ub’ denote the sizes of certain ubiquitin-related proteins. ‘UBB⁺¹’ indicates a ubiquitin B variant, while ‘⁴⁸Ub’ refers to a polyubiquitin chain linked via its K48. ‘Ub_n-UBB⁺¹’ represents a polyubiquitinated UBB⁺¹ variant (trimer, tetramer, and subsequent oligomeric states), and ‘-Ub-UBB⁺¹’ indicates a single ubiquitin addition to UBB⁺¹ (dimer).

The results display a control group (K) with numbers 005 to 006 (on the left side) and an experimental group (S) marked with numbers from 112 to 116. On the right side, bands of synthesized Ub-⁴⁸UBB⁺¹ dimer are visible. A lack of bands in the Ub^{WT} (wild type) section was also observed, indicating antibody binding specificity exclusively to the modified form. The WB analysis included proteins from the peripheral blood of the control group and patients diagnosed with schizophrenia. Detection using anti-Ub⁺¹ antibodies (40B3) indicates the predominant presence of polymeric forms of Ub_n-UBB⁺¹ in both schizophrenia patients (S) and control samples (K).

On Figure 1, a smear between 25 kDa and 250 kDa corresponds to polychains of Ub_n-UBB⁺¹, 15 kDa to Ub_n-UBB⁺¹ dimers, indicating the expression of Ub-UBB⁺¹ conjugates. This analysis, part of a broader research process, includes the synthesis of Ub-⁴⁸UBB⁺¹ dimers, with bands indicating specific proteins and their modifications, which is significant for understanding biological processes such as protein degradation or DNA repair. Additional labels on the bands, such as ‘Ub^{WT}’, ‘UBB⁺¹’, ‘Ub-⁴⁸UBB⁺¹’, and ‘Ub_n-UBB⁺¹’, indicate various forms of Ub, with a scale on the left side showing measurements in kilodaltons (kDa) for assessing protein size.

3.2. Results of the AI Models Analysis

Comparing the responses from four AI models (Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4) on the WB analysis, we can discern similarities and differences in their interpretative capabilities (Figure 2). The entire conversation is available in File S1 in Supplementary Materials.

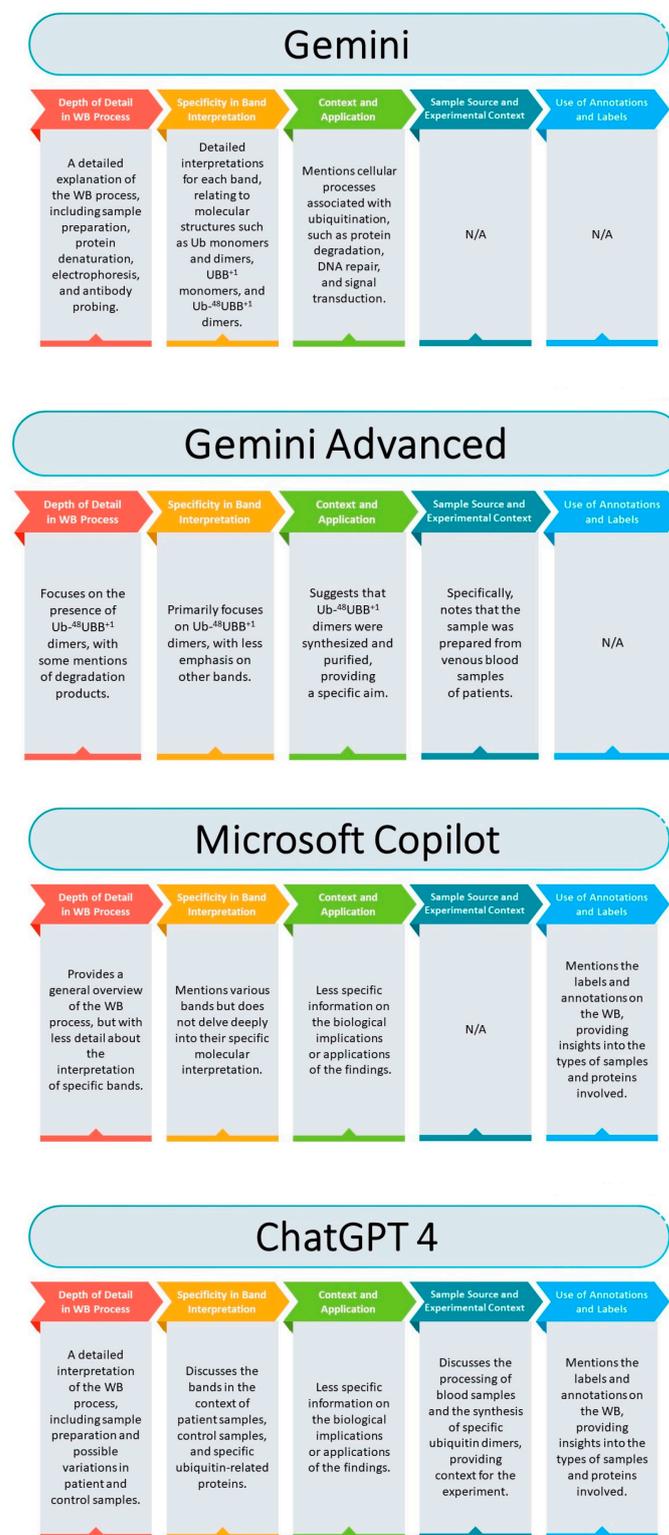


Figure 2. Comparison of the interpretative abilities of four AI models (Gemini, Gemini Advanced, Microsoft Copilot, ChatGPT 4) in the analysis of WB images containing UBB in the context of

schizophrenia research. The figure details the depth of description of the WB process, the specificity of the interpretation of individual bands, the context and potential applications of the findings, the source of the sample and experimental context, and the use of annotations and labels by each model. The comparison highlights differences in each model's approach to biological data analysis, emphasizing their unique features and potential in the field of biomedical research. N/A—no data.

In summary, while all models effectively address the subject of WB analysis and recognize the key elements of the process, there are variations in the depth of their explanations, the specificity of band interpretations, and the context provided for the experimental setup. Gemini provides a comprehensive explanation with detailed band interpretation, Gemini Advanced focuses on specific dimers, Microsoft Copilot gives a general overview, and ChatGPT 4 offers detailed interpretations with additional context on sample types and experimental design.

4. Discussion

In our study, we analyzed the responses of four AI models—Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4—regarding the analysis of WB imagery, affirming the posited hypothesis. Each AI model demonstrated the ability to conduct fundamental basic analysis and interpretation of WB images, identifying key elements such as the presence and characteristics of individual protein bands and describing the fundamental steps of the WB procedure. Gemini distinguished itself with a detailed description of the WB process and band interpretation, emphasizing their biological significance. Gemini Advanced focused on identifying and interpreting specific bands, particularly highlighting Ub-⁴⁸UBB⁺¹ dimers. Microsoft Copilot provided a general overview of the WB image, noting key sections and bands, but with less technical depth. ChatGPT 4, on the other hand, offered a detailed interpretation of bands in the context of patient samples and standards, encompassing both biological and technical aspects. The differences in detail and context among the models were consistent with the assumptions of the hypothesis, allowing it to be considered confirmed.

Our findings indicate significant differences in the manner of WB data interpretation by different AI models, which significantly impacts their utility in biomedical research. Analyzing the responses of the four AI models, notable differences were observed in the depth of details, specificity of band interpretation, and the context and application of the provided information.

A fundamental observation is that all models effectively recognize and describe WB analysis, focusing on identifying specific protein bands. This basic competency in recognizing key elements of the WB process aligns with the general understanding of molecular biology techniques trained in AI models. However, differences in the level of accuracy and depth of interpretation among the models underscore that not all AI tools are equally effective in delivering the detailed and contextual information necessary for advanced scientific analysis.

The Gemini model stands out with a detailed description of the WB process and precise band interpretation, which can be useful in detailed biological and biochemical studies. This model appears to be more useful in contexts where a deep technical analysis and understanding of molecular processes are required.

In contrast, Gemini Advanced, focusing on the presence of Ub-⁴⁸UBB⁺¹ dimers, demonstrates the ability to identify specific elements in samples, which can be useful in more targeted studies, such as research on specific protein modifications.

Microsoft Copilot, providing a general overview of the WB process, can be helpful in educational situations or preliminary analyses, where detailed technical knowledge is not crucial but a general understanding of the process is still needed.

ChatGPT 4, presenting detailed interpretations of bands in the context of patient samples and control standards, is particularly useful in clinical and diagnostic research, where understanding the relationship between bands and specific biological states is key.

These differences can be attributed to the specific training and architecture of each model. Each model has been trained and optimized in different contexts, influencing its ability to process and interpret specific scientific data. For example, models trained on a broad range of scientific literature may perform better with general interpretations, while specialized models may excel in identifying and analyzing detailed technical aspects.

In summary, our results indicate that while each AI model can be useful in WB data analysis, their effectiveness and utility depend on the research context. The choice of AI tool should be dictated by the specific needs of the study, domain knowledge, and requirements for technical and contextual depth of analysis.

In the realm of proteomics, the integration of AI has shown significant advancements. For instance, Mann et al. [17] demonstrated how machine learning, particularly deep learning, is now capable of predicting experimental peptide measurements from amino acid sequences alone, a breakthrough that could dramatically improve the quality and reliability of analytical workflows in proteomics. Similarly, Vishnoi et al. [18] explored the use of AI and machine learning for protein toxicity prediction using proteomics data, further demonstrating the versatility of AI in handling complex biological datasets. Additionally, the study by Cui et al. [19] focusing on protein–DNA/RNA interactions and machine intelligence tools, reveals the expanding scope of AI applications in understanding intricate molecular interactions. The escalating challenges in ensuring scientific integrity, particularly concerning the authenticity of WB images, have been thoroughly examined in recent studies by Qiet al. and Mandelli et al., both of which underscore the sophisticated threats posed by digital manipulations [20,21]. Qiet al. revealed the efficacy of generative adversarial nets (GANs) in creating WB images so convincing that they are virtually indistinguishable from actual specimens. This discovery poses a profound challenge to traditional fraud detection methodologies, which primarily rely on visual inspections and pattern recognition but yield an accuracy barely above that of a blind guess. The research further demonstrated that detection accuracy does not significantly correlate with academic qualifications but shows a modest improvement with increased experience in WB-related research [20].

On a complementary note, Mandelli et al. tackled the issue of WB image falsification by generating a comprehensive dataset containing over 14,000 real and 24,000 synthetic images through the use of GANs and denoising diffusion probabilistic models (DDPMs). Their findings corroborate the difficulty in distinguishing between authentic and counterfeit images, emphasizing that tools trained on genuine images can, indeed, detect forgeries. Moreover, their study highlights the resilience of these detection techniques against common image alterations such as compression, although challenges remain in accurately identifying images modified through scaling. However, synthetic images crafted using DDPMs were reliably identified post-resizing, marking a critical advancement in the protection of research integrity [21].

To address these emerging threats, Qi et al. recommend the implementation of stricter verification measures, including the mandatory submission of WB images alongside a unique identifier generated by laboratory equipment and the peer review of these images in conjunction with article submissions. Such measures are proposed to bolster the verification process and mitigate the risks of scientific fraud, ensuring that the WB images presented in scholarly articles are authentic and trustworthy. Together, these studies illuminate the path forward in combating digital falsification, advocating for a combination of advanced detection technologies and rigorous review protocols to uphold the standards of scientific integrity [20].

From a broader perspective, our study highlights the potential of AI in complementing and enhancing research in molecular biology. The ability of AI models like Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4 to provide precise analyses in specialized scientific fields without specific training is a testament to the evolving versatility of AI technology. This opens doors to numerous applications where AI can assist in the interpretation of complex biological data, thereby accelerating research and discoveries in fields such as molecular biology, genomics, and proteomics.

Furthermore, the successful application of these models suggests a future where AI could be specifically trained or fine-tuned using datasets from specialized fields such as WB analysis. Such advancements could lead to even more precise and nuanced interpretations, potentially revolutionizing the way scientists analyze experimental data.

In conclusion, the confirmation of our hypothesis not only demonstrates the advanced capabilities of AI models like ChatGPT 4 but also signals a promising future for AI applications in scientific research, particularly in fields requiring the interpretation of complex data sets. The potential of AI to significantly contribute at various stages of scientific research, from designing experiments to analyzing and interpreting data, is immense and largely untapped.

Although this study provides valuable insights into the interpretative abilities of different AI models in the context of WB analysis, its results should be interpreted with consideration of the following limitations:

1. **AI Model Scope:** This study was limited to analyzing responses from only four AI models (Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4), which does not encompass the full spectrum of available AI tools. There are other AI models that might offer different perspectives or unique interpretative abilities, potentially altering or broadening our understanding of AI's capability in WB data analysis.
2. **Training Data Specificity:** AI models are trained on data that may not fully represent all aspects of WB analysis. The limitations and scope of the training data directly impact the quality and accuracy of the models' generated responses. Therefore, responses might reflect the knowledge and limitations specific to the data on which the models were trained.
3. **WB Analysis Complexity:** WB is a complex technique with many variables, including different sample types, detection methods, and antibody specificities. AI models may not be fully equipped to handle all the nuances and complexities associated with WB analysis, particularly in unusual or complicated cases.
4. **Contextual Interpretation:** While AI models may effectively identify and interpret basic elements of WB analysis, their ability to understand deeper biological or clinical context is limited. Interpretation of WB results in the context of specific diseases, disorders, or molecular mechanisms may require human judgment and specialized knowledge.
5. **AI Development Dynamics:** AI technology is rapidly evolving, meaning that findings and conclusions could quickly become outdated. New models and updates to existing tools might offer improved accuracy, deeper analysis, and new interpretive capabilities.

Subjectivity in Interpretation: AI model responses may be somewhat subjective, depending on how the question was formulated, which could affect the analysis results. The human-like approach to interpretation and question formulation can also influence the nature of responses.

Potential Biases

Selection Bias: The choice of AI models and their underlying training datasets might introduce biases toward certain interpretations or overlook specific aspects of the WB images.

Interpretation Bias: The nature of AI-generated interpretations can be influenced by the input format and the phrasing of questions posed to the models, which might skew the analysis.

Recognizing these biases and limitations is crucial for the appropriate application of AI in scientific research and emphasizes the need for continuous oversight and validation by human experts. Future studies should aim to address these limitations by employing a more diverse set of AI models, developing more specialized training datasets, and incorporating rigorous validation processes to assess the accuracy of AI interpretations.

5. Conclusions

The comparative study of various AI models, such as Gemini, Gemini Advanced, Microsoft Copilot, and ChatGPT 4, has demonstrated that each possesses a unique approach

and interpretative capabilities, particularly in the context of WB data analysis. Their diversity in terms of detail, specificity, and contextual understanding underscores the need to tailor the choice of AI tool to the specific requirements of the research. These models have the potential to significantly support biomedical research, offering rapid and efficient assistance in identifying key elements such as protein bands. However, their limitations suggest the necessity of integrating them with human expertise, especially in terms of biological understanding and result evaluation. Furthermore, AI models can serve as educational tools in teaching WB analysis techniques, proving useful for students and emerging scientists. This study sheds light on the need for further research into the utilization and optimization of AI tools in molecular biology and other scientific fields. Given the dynamic evolution of AI technology, it is crucial to monitor new updates and models that may offer enhanced interpretative and analytical capabilities. Future research should explore integrating these AI tools into real-world biomedical applications, such as diagnostic processes, therapeutic development, and personalized medicine. Moreover, the development of AI models trained specifically on WB and other proteomic data could open new avenues for more precise and accurate data interpretation. This approach could lead to breakthroughs in understanding diseases at a molecular level and accelerating drug discovery by identifying potential targets more efficiently.

To further enhance the reliability and applicability of AI in scientific research, future studies could also focus on developing hybrid models that combine the strengths of various AI tools. This could potentially lead to the creation of a superlative model that leverages the detailed analytical capabilities of Gemini, the contextual understanding of ChatGPT 4, and the broad data integration features of Microsoft Copilot.

Such efforts could lead to significant breakthroughs in our understanding of complex biological systems and the development of innovative solutions to pressing health challenges.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app14104149/s1>, File S1: AI Interpretation of UBB⁺¹ in the Western Blot Image.

Author Contributions: Conceptualization, A.F. and M.C.; methodology, M.C., A.F., A.Z.-F., M.P. and B.P.; software, R.F.; investigation, M.P., A.F., B.P. and M.C.; data curation, A.F., M.C. and M.P.; writing—original draft preparation, A.F., M.C., A.Z.-F., R.F., M.P. and B.P.; writing—review and editing, M.C., A.F., B.P., A.Z.-F., M.P., E.N. and K.Z.; supervision, M.C., B.P., E.N. and K.Z.; funding acquisition, A.F., B.P., E.N. and K.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The entire procedure was approved by the Bioethics Committee of the Medical University of Lodz, in accordance with decision number RNN/208/17/KE.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: Many thanks are extended to Agnieszka Strzała for linguistic proofreading and to Robert Fabijan for substantive support in the field of AI and for assistance in designing the present study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ovchinnikova, K.; Kovalev, V.; Stuart, L.; Alexandrov, T. OffsampleAI: Artificial intelligence approach to recognize off-sample mass spectrometry images. *BMC Bioinform.* **2020**, *21*, 129. [CrossRef]
2. Emmert-Streib, F. Grand Challenges for Artificial Intelligence in Molecular Medicine. *Front. Mol. Med.* **2021**, *1*, 734659. [CrossRef]

3. Jonak, K.; Jędrasiak, K.; Polański, A.; Puszyński, K. Application of Image Processing Algorithms in Proteomics: Automatic Analysis of 2-D Gel Electrophoresis Images from Western Blot Assay. In *Computer Vision and Graphics*; Bolc, L., Tadeusiewicz, R., Chmielewski, L.J., Wojciechowski, K., Eds.; ICCVG 2012. Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2012; pp. 433–440. [[CrossRef](#)]
4. Bhalla, S.; Laganà, A. Artificial Intelligence for Precision Oncology. *Adv. Exp. Med. Biol.* **2022**, *1361*, 249–268. [[CrossRef](#)] [[PubMed](#)]
5. Sandbrink, J.B. Artificial intelligence and biological misuse: Differentiating risks of language models and biological design tools. *arXiv* **2023**, arXiv:2306.13952. [[CrossRef](#)]
6. Waisberg, E.; Ong, J.; Masalkhi, M.; Zaman, N.; Sarker, P.; Lee, A.G.; Tavakkoli, A. GPT-4 and medical image analysis: Strengths, weaknesses and future directions. *J. Med. Artif. Intell.* **2023**, *6*, 29. [[CrossRef](#)]
7. Microsoft Copilot. Available online: <https://www.theverge.com/2023/12/19/24008279/microsoft-copilot-suno-ai-music-generator-extension> (accessed on 17 February 2024).
8. Cortana. Available online: <https://support.microsoft.com/en-us/topic/end-of-support-for-cortana-d025b39f-ee5b-4836-a954-0ab646ee1efa> (accessed on 17 February 2024).
9. Google Gemini AI. The Future of Artificial Intelligence. Available online: <https://digitalfloats.com/google-gemini-the-future-of-artificial-intelligence/> (accessed on 17 February 2024).
10. Gemini. Available online: https://gemini.google.com/advanced?utm_source=deepmind&utm_medium=owned&utm_campaign=gdm-site_learn (accessed on 17 February 2024).
11. Chojnacki, M.; Zhang, D.; Talarowska, M.; Gałeczki, P.; Szemraj, J.; Fushman, D.; Nakasone, M.A. Characterizing polyubiquitinated forms of the neurodegenerative ubiquitin mutant UBB⁺¹. *FEBS Lett.* **2016**, *590*, 4573–4585. [[CrossRef](#)] [[PubMed](#)]
12. Rubio, M.D.; Wood, K.; Haroutunian, V.; Meador-Woodruff, J.H. Dysfunction of the ubiquitin proteasome and ubiquitin-like systems in schizophrenia. *Neuropsychopharmacology* **2013**, *38*, 1910–1920. [[CrossRef](#)] [[PubMed](#)]
13. Dong, K.C.; Helgason, E.; Yu, C.; Phu, L.; Arnott, D.P.; Bosanac, I.; Compaan, D.M.; Huang, O.W.; Fedorova, A.V.; Kirkpatrick, D.S.; et al. Preparation of distinct ubiquitin chain reagents of high purity and yield. *Structure* **2011**, *19*, 1053–1063. [[CrossRef](#)] [[PubMed](#)]
14. Bhardwaz, S.; Kumar, J. An Extensive Comparative Analysis of Chatbot Technologies—ChatGPT, Google BARD and Microsoft Bing. In Proceedings of the 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 4–6 May 2023; pp. 673–679. [[CrossRef](#)]
15. Brin, D.; Sorin, V.; Barash, Y.; Konen, E.; Glicksberg, B.S.; Nadkarni, G.; Klang, E. Assessing GPT-4 Multimodal Performance in Radiological Image Analysis. *medRxiv* **2023**. medRxiv:2023.11.15.23298583.
16. Carlà, M.M.; Gambini, G.; Baldascino, A.; Giannuzzi, F.; Boselli, F.; Crincoli, E.; D’Onofrio, N.C.; Rizzo, S. Exploring AI-chatbots’ capability to suggest surgical planning in ophthalmology: ChatGPT versus Google Gemini analysis of retinal detachment cases. *Br. J. Ophthalmol.* **2024**. [[CrossRef](#)]
17. Mann, M.; Kumar, C.; Zeng, W.F.; Strauss, M.T. Artificial intelligence for proteomics and biomarker discovery. *Cell Syst.* **2021**, *12*, 759–770. [[CrossRef](#)] [[PubMed](#)]
18. Vishnoi, S.; Matre, H.; Garg, P.; Pandey, S.K. Artificial intelligence and machine learning for protein toxicity prediction using proteomics data. *Chem. Biol. Drug. Des.* **2020**, *96*, 902–920. [[CrossRef](#)] [[PubMed](#)]
19. Cui, F.; Zhang, Z.; Cao, C.; Zou, Q.; Chen, D.; Su, X. Protein-DNA/RNA interactions: Machine intelligence tools and approaches in the era of artificial intelligence and big data. *Proteomics* **2022**, *22*, e2100197. [[CrossRef](#)] [[PubMed](#)]
20. Qi, C.; Zhang, J.; Luo, P. Emerging Concern of Scientific Fraud: Deep Learning and Image Manipulation. *bioRxiv* **2020**. bioRxiv:2020.11.24.395319.
21. Mandelli, S.; Cozzolino, D.; Cannas, E.D.; Cardenuto, J.P.; Moreira, D.; Bestagini, P.; Scheirer, W.J.; Rocha, A.; Verdoliva, L.; Tubaro, S.; et al. Forensic analysis of synthetically generated western blot images. *IEEE Access* **2022**, *10*, 59919–59932. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.