

# A simple workflow to identify novel small linear motif (SLiM)-mediated interactions with AlphaFold

Martin Veinstein<sup>1</sup>, Victor Janssens<sup>1</sup>, Bogdan I. Iorga<sup>2</sup>, Raphaël Helaers<sup>1</sup>, Thomas Michiels<sup>1,\*,†</sup>, Frederic Sorgeloos<sup>3,†</sup>

<sup>1</sup>Université Catholique de Louvain, de Duve Institute, Brussels, Belgium

<sup>2</sup>Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles (ICSN), Gif-sur-Yvette, France

<sup>3</sup>INRS-Centre Armand-Frappier Santé Biotechnologie, Laval, Québec, Canada

\*Corresponding author. Université Catholique de Louvain, de Duve Institute, 74 (VIRO - B1.74.07) avenue Hippocrate, B-1200 Brussels, Belgium.

E-mail: [thomas.michiels@uclouvain.be](mailto:thomas.michiels@uclouvain.be)

†Thomas Michiels and Frederic Sorgeloos co-last authors.

## Abstract

Short linear motifs (SLiMs) are highly compact interaction modules embedded within disordered protein regions and are increasingly recognized for their central role in maintaining cellular homeostasis. Due to their small size, degeneracy and transient binding, SLiMs remain difficult to detect both experimentally and computationally. Here, we show that AlphaFold (AF), used via ColabFold, offers a practical and accessible alternative for *in-silico* screening of new SLiMs targeting a protein of interest. Unlike previous studies that evaluated AlphaFold2 (AF2) using structure-derived benchmarks, we extend this by assessing both AF2 and AF3, using a structure-independent benchmark of 26 interactions absent from PDB homology, and showing that MiniPAE is the most suited AlphaFold metric for SLiM screening. We also generated an unbalanced dataset with a large excess of non-binders mimicking real-world blind screening, revealing a critical limitation in AlphaFold's specificity for SLiM detection. To circumvent this constraint, we propose both a SLiM screening strategy and an adaptative scoring threshold. For greater accessibility, we provide a streamlined and cost-effective AF analysis workflow requiring no local installation or computation. To overcome challenges associated with SLiM validation, we also introduce a highly sensitive detection method based on proximity labeling in living cells. This workflow was used to identify and experimentally validate 13 new SLiMs that mediate binding to ribosomal protein S6 kinase A3 (RPS6KA3 or RSK2). By leveraging ColabFold and MiniPAE available through Colab notebooks, our approach provides a scalable and widely accessible strategy for identifying functional SLiMs in proteins of interest. MiniPAE can be accessed at <https://github.com/martinovein/MiniPAE>

**Keywords:** small linear motif; SLiM; AlphaFold; intrinsically disordered protein; protein–protein interactions; RSK; p90 ribosomal S6 kinase

## Introduction

Short linear motifs (SLiMs) are short regions of 3–10 amino acids embedded within intrinsically disordered regions (IDRs) that mediate protein–protein interactions (PPIs). These interactions generally occur between a SLiM and a structured protein domain and exhibit low-to-mid micromolar affinities, making them well suited for dynamic and reversible interactions that regulate cell signaling [1]. SLiMs control a wide range of biological processes and are catalogued in resources such as the eukaryotic linear motif (ELM) database, which currently contains around 1500 annotated human SLiMs [2]. However, given the estimated tens of thousands of human SLiMs, the vast majority of them remain undiscovered [1].

Despite their importance, SLiMs are notoriously challenging to detect directly due to their small size and weak binding affinities. Alternative *in-vitro* methods have been developed to improve detection sensitivity [3]. However, these approaches remain labor-intensive, difficult to scale-up and still face intrinsic affinity limitations. As such, computational approaches have emerged as powerful complementary tools to accelerate SLiM discovery.

Despite not being specifically designed for this purpose, AlphaFold2 (AF2) has shown an unexpected ability to model SLiM-mediated interactions. An early study demonstrated that AF2 could accurately predict the structure of interactions between IDRs and their structured binding domains, even for protein pairs entirely absent from its training set [4].

Subsequent benchmarking by Lee *et al.* provided a systematic comparison of different AAF-based metrics for SLiM–domain interactions [5]. Notably, they were among the first to apply AF at scale to screen for SLiM-mediated PPIs across an entire signaling network. While their results highlighted the limited specificity of AF in this context, they did not propose a strategy to address this limitation.

More recently, Omid *et al.* conducted a comprehensive analysis of AF2's ability to model a larger variety of IDR-mediated interactions including SLiMs, molecular recognition features (MoRFs), among others [6]. Importantly, they introduced two new residue-level metrics: MinD, which evaluates the minimum distance between residues across interfaces, and MiniPAE, derived from AF's predicted aligned error (PAE) maps, to identify potential interaction regions within full-length IDRs.

Received: June 6, 2025. Revised: August 1, 2025. Accepted: August 19, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, we build on previous evaluation of AF2 for SLiM discovery by extending the comparison to AF3 [7] and introducing a structure-independent benchmark of 26 interactions lacking PDB homology. Using a large, unbalanced dataset mimicking real-world blind screening, we expose a critical limitation in AF's specificity. To address this, we propose a new SLiM screening strategy that incorporates an adaptive scoring threshold to select high-confidence predictions. This analysis led to a streamlined and accessible workflow for SLiM discovery, which we have made available through a user-friendly Colab notebook. We illustrated the power of this approach with a case study focused on identifying RSK-binding SLiMs. To complement our *in-silico* predictions and address the challenge of low-affinity interactions, we further introduced a sensitive proximity labeling method, enabling functional validation, in cells, of candidate SLiMs predicted by our pipeline.

## Results

### Benchmark SLiM datasets

To assess AF's capacity to identify SLiMs within binary interaction predictions, we constructed a benchmark dataset derived from the ELM database, a manually curated repository of experimentally validated SLiMs. Given the high redundancy of SLiMs, the ELM database groups analogous motifs under a common classification, termed ELM\_Class. Among the 161 ELM\_Classes annotated in humans, we selected 121 representative instances, each corresponding to a distinct ELM\_Class, to serve as our positive dataset, hereafter referred to as the ELM\_dataset (see Methods, Fig. 1A, Supplementary data S1).

Unlike structure-centric benchmarks, our evaluation does not require the existence of homologous complexes in the protein data bank (PDB). This allowed us to form a subdataset of 26 SLiM-partner pairs which have no available structural homologs. This dataset was named noPDB\_dataset (Fig. 1A, Supplementary data S1).

To rigorously evaluate AF's predictive performance in different screening scenarios, we generated a large negative control set by randomly shuffling the binding partners from our positive dataset, following the strategy of Lee et al. [5] (Fig. 1B, Supplementary Data S1). From all possible shuffled pairs, 1100 likely non-binding combinations were randomly selected to serve as 'non-binders' (see Methods).

### Comparison of different AF2/3 metrics to discriminate SLiM-binders from non-binders

AF's utility in SLiM discovery can be conceptually divided into two challenges. The first is binary 'classification'—determining whether a SLiM mediates binding to a partner at all. The second is 'localization'—identifying the specific region that functions as a SLiM. For those tasks different AF-based metrics can be used.

Among those, pLDDT has garnered considerable attention as a metric for localizing SLiMs [8]. However, it suffers from a fundamental limitation: its scores tend to increase not only where a disordered region folds upon binding (the signal of interest), but also in pre-structured regions that are not involved in the interaction (noise). To address this limitation, we developed AlphaSLiM, a new scoring metric that quantifies the change in per-residue pLDDT between the unbound and bound states of an IDR (see Methods). We compared AlphaSLiM to four other AF-based metrics: the model confidence (MC) computed by AF, the average\_model proposed by Schmid et al. [9], and the minD and miniPAE scores proposed by Omid et al. [6].

Using our ELM\_dataset, we first evaluated these metrics for their ability to distinguish true SLiM-binders from non-binders. All metrics—AlphaSLiM, MC, average\_model, minD, and miniPAE—achieved comparable performance, with ROC area under the curve (AUC) values around 70% ( $\pm 3\%$ ) (Supplementary Fig. S1A, Supplementary data S1) suggesting that AF-based scoring functions can moderately discriminate binders from non-binders. Importantly, when we repeated the evaluation using the noPDB\_dataset, results remained consistent (AUC = 70%  $\pm$  4%, Fig. 1B, Supplementary data S1), indicating that the observed performance of AF-based metrics in distinguishing binders from non-binders did not arise from structural biases linked to PDB templates.

To assess AF's performance under real world SLiM screening conditions in which non-binders vastly outnumber true binders, we leveraged our large nonBinder\_dataset to generate precision-recall curves using increasingly unbalanced datasets (1:1, 1:5, and 1:25 binder to non-binder ratios). As shown in Fig. 1C, AF's performance dropped sharply as the ratio became more unbalanced. The AUC fell from  $\sim 70\%$  in a balanced setting (1:1) to 37% and 14% at 1:5 and 1:25 ratios, respectively. These results highlight AF's limited specificity in blind screening contexts when true binders are rare (e.g. 1:1000).

To evaluate whether newer model architectures improve SLiM detection, we compared AF2 and AF3. To avoid biases stemming from differences in training data, especially with respect to known structural templates, we restricted this comparison to the noPDB\_dataset. Our results revealed comparable performance across AF2 and AlphaFold3 metrics in terms of binder versus non-binder discrimination, suggesting that despite its architectural advances, AF3 MC does not provide a clear advantage for SLiM binders from non-binders discrimination (Fig. 1B, Supplementary data S1).

### Comparison of different AF2/3 metrics to identify new SLiMs within full-length proteins

We next assessed AF's ability to localize a binding SLiM within full-length proteins. For this, we compared the per-residue score within the SLiM region when paired with its true partner to the score within the same region when paired with a random non-binder (Fig. 2A). With our ELM\_dataset, all three AF-derived metrics showed similar overall performances (AUC  $\sim 78\%$ , Fig. 2B), but miniPAE outperformed the others at FDR = 0. This advantage disappeared with the noPDB\_dataset (Fig. 2C and D; Supplementary Fig. S1B), suggesting that miniPAE's precision is driven by structural similarity to known templates. Of note, physics-based metric like the buried surface area (BSA) metric performed significantly worse, with recall at FDR = 0% falling from  $\sim 30\%$  to 4% (Fig. 2B).

The large overlap between true/false motif calls from AlphaSLiM, minD, and miniPAE (Supplementary Fig. S1C) suggests that combining these metrics offers little added benefit. Additionally, although shorter prediction lengths are known to improve structural accuracy [4, 5], we found no correlation between prediction length and motif detection (Spearman  $\rho = 0.03$  and  $-0.17$  for ELM\_dataset and noPDB\_dataset, respectively; Supplementary Fig. S1D-E).

Since both AF2 and AF3 output PAE scores, we used the noPDB\_dataset to compare their ability to localize SLiMs within proteins. AF3 showed slightly lower performance than AF2 in this task (Fig. 2C, Supplementary Fig. S1B; Supplementary data S1).

Figure 1C highlighted a sharp drop in AF2 performance under unbalanced conditions where non-binders greatly outnumber binders, a common scenario in SLiM screening. To explore this

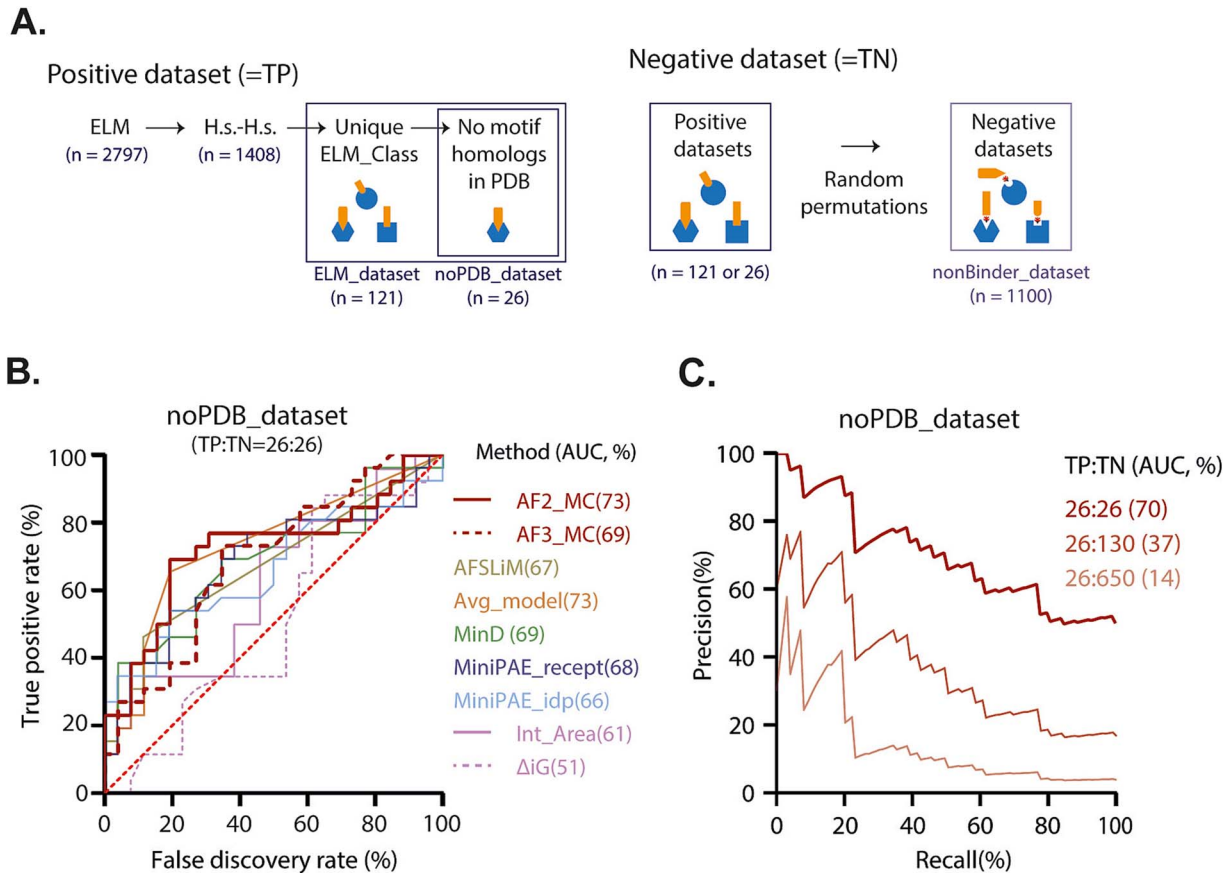


Figure 1. Benchmarking of AlphaFold-based metrics for SLiM-mediated interaction detection. (A) Construction of benchmark datasets. A non-redundant set of 121 human SLiM-partner pairs, each representing a unique ELM\_Class, was extracted from the ELM database resulting in the ELM\_dataset (Supplementary data S1). A subset of 26 instances lacking structural homologs in the protein data bank (PDB) constituted the noPDB\_dataset. Negative datasets were generated by randomly permuting the interaction partners within each positive dataset (see method). (B) Receiver operating characteristic (ROC) curves comparing true-binders to randomly shuffled non-binders, using the balanced noPDB\_dataset. Seven scoring metrics are compared: Model confidence (MC), AlphaSLiM (AFSLiM), average\_model from Schmid *et al.* [9] (Avg\_model), the minD and miniPAE metrics from Omidi *et al.* [6], interface area, and solvation free energy ( $\Delta iG$ ). Area under the curve (AUC) values are indicated between parentheses for each metric. AF2 = AlphaFold2, AF3 = AlphaFold3, TP = true positives, and TN = true negatives. (C) Precision-recall curves evaluating AlphaFold2 Model confidence ability to distinguish binders from non-binders at different binder to non-binders ratios.

further, we leveraged our larger nonBinder\_dataset ( $n = 1100$ ) to evaluate, using precision-recall curves, the ability of AF2's MiniPAE metric to identify SLiMs within full-length proteins in increasingly unbalanced scenarios. Figure 2E confirmed a marked decline in performance as the binder-to-non-binder ratio increased, with AUC values dropping from 70% (1:1) to 43% (1:5) and 27% (1:25).

To address this limitation, we analyzed how the distribution of MiniPAE scores shifts with increasing numbers of non-binders. This analysis, presented in Fig. 2F, allowed us to derive an adaptive thresholding scheme by fitting a power-law function to the data. This formula enables dynamic adjustment of the scoring threshold based on the scale of the screening campaign: when screening hundreds of potential partners, a stringent threshold minimizes false positives, while in focused analyses involving a known interactor, a less stringent threshold maximizes recovery.

### AF-based workflow to identify new SLiMs binding a protein of interest

Building on our benchmarking results, we propose a step-by-step workflow to discover new SLiMs binding a specific protein (Fig. 3), while highlighting common pitfalls in such AF screenings. The first critical decision lies in selecting candidate protein partners

for binary prediction. An overambitious approach would be to perform blind screening against entire proteomes. In line with this, Figs 1C and 2E demonstrate that AF-based metrics perform poorly when non-binders vastly outnumber true-binders, a behavior also reported by Schmid *et al.* [9]. To mitigate this, we recommend selecting partners from curated interaction databases such as BioGRID [10] or IntAct [11], or from in-house datasets. Given that AF2 slightly outperforms AF3, and is readily accessible, we suggest its use for most cases.

To streamline SLiM retrieval, we adapted the MiniPAE script from Omidi *et al.* [6] to analyze AF2 output files, returning MiniPAE scores, motif sequences, and coordinates. Since experimental validation is often more limiting than prediction time, a stringent threshold is recommended to prioritize high-confidence candidates. In line, an adaptive MiniPAE threshold was defined in Fig. 2F as:

$$\text{MiniPAE threshold} = 30.29 \times n^{-0.2803} - 3.82$$

where  $n$  represents the number of predictions analysed.

With regard to SLiMs, validating them remains a significant challenge given the low affinity interaction that they often mediate [3]. To mitigate this, we propose a proximity labeling assay as a low-cost, sensitive method to detect interactions in living cells.

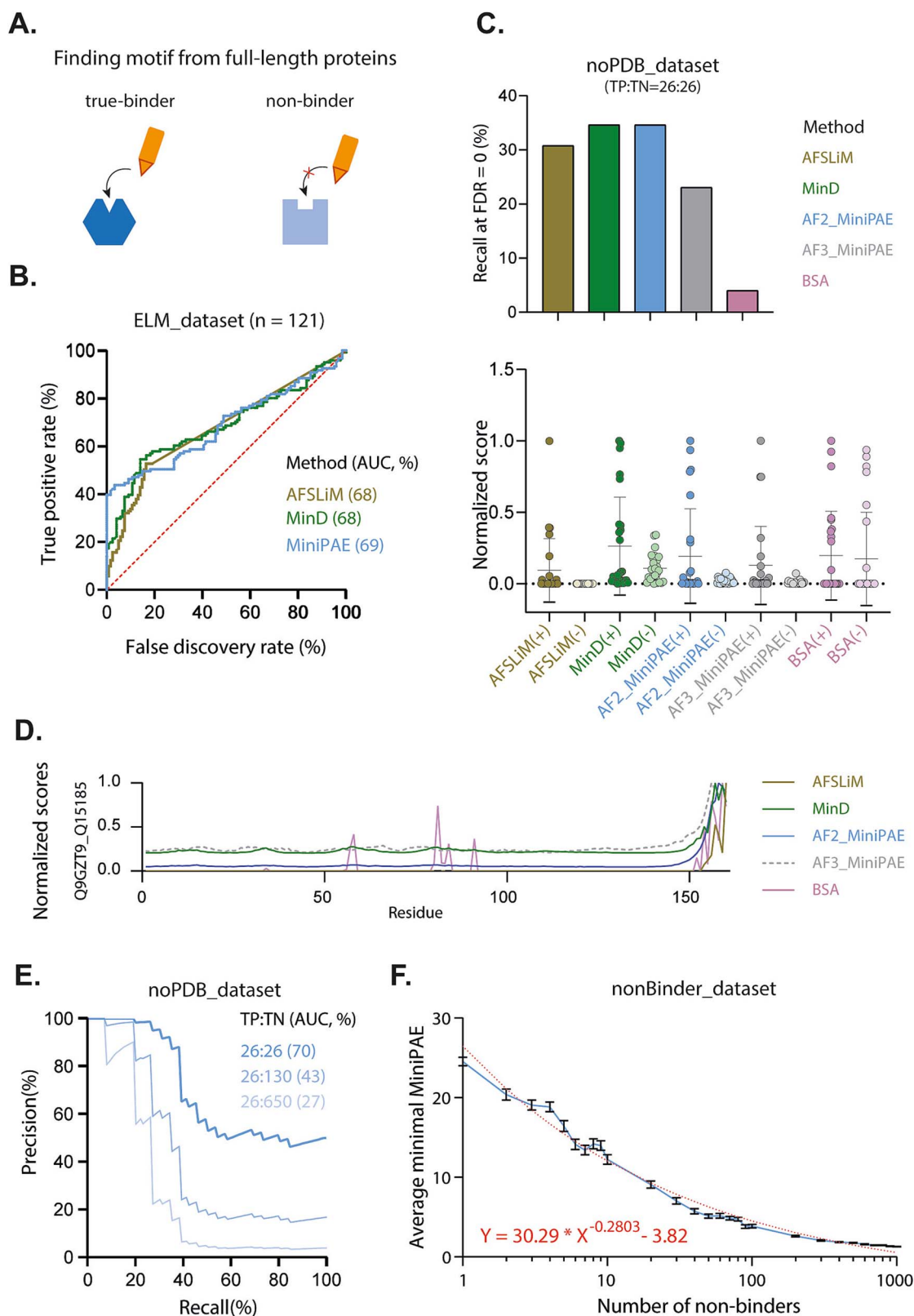


Figure 2. Using AlphaFold metrics to localize SLiMs within full-length proteins. (A) Schematic of the analysis: Scores within the motif region when paired with its true-binder are compared to the score within the same region when paired with a random non-binder. (B) ROC curves for the ELM\_dataset. (C) Top: True positive rate (TPR, %) at FDR=0. Bottom: Distribution of scores in the motif region for true-binders (+) versus non-binders (-) on a 0,1 scale (see method). BSA = buried surface area. (D) Example of per-residue score profile. (E) Precision-recall curves evaluating AlphaFold2 MiniPAE ability to identify SLiM with full-length proteins in unbalanced scenarios. TP:TN indicates binder-to-non-binder ratios. (F) Distribution of the average minimal value for MiniPAE score with increasing numbers of non-binders (see method). A power-law curve was fitted to model the relationship between dataset size and optimal threshold.

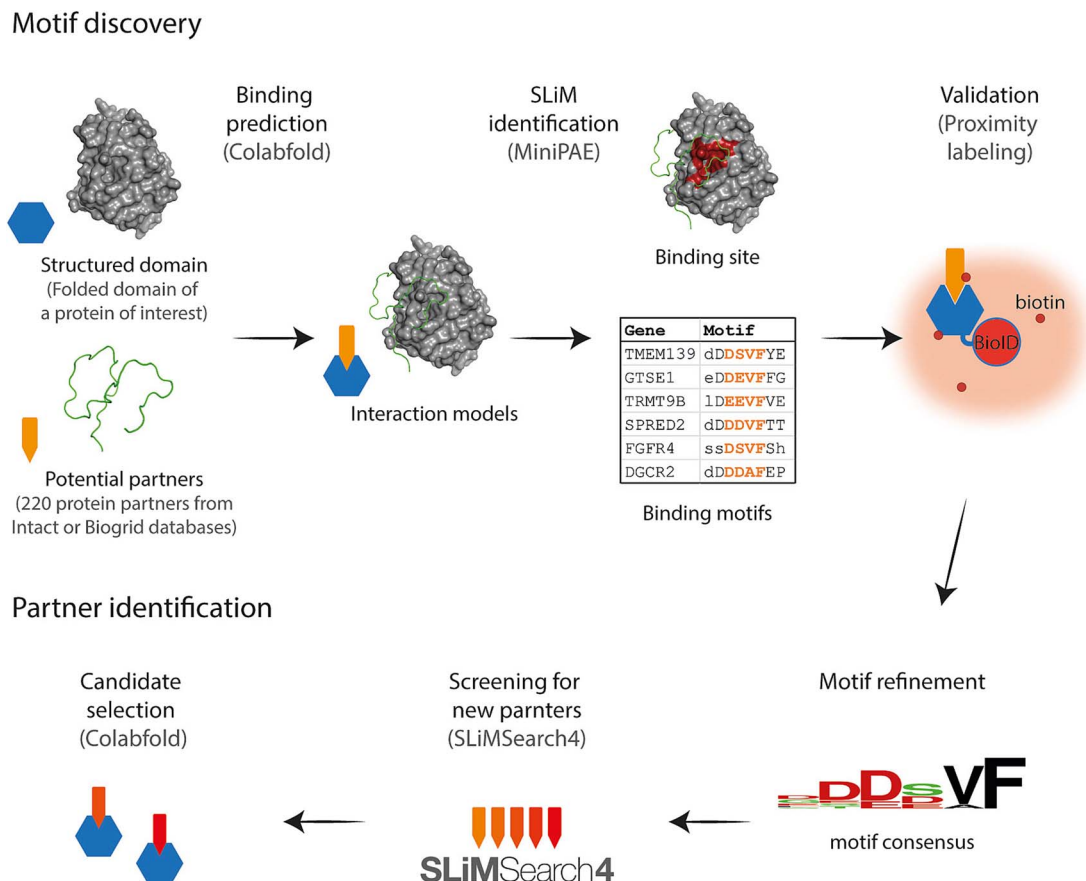


Figure 3. An AlphaFold-based workflow to identify new SLiMs. ColabFold is used to perform binary binding predictions between a protein (or domain) of interest and selected candidates from interaction databases or in-house datasets. The MiniPAE script extracts high-confidence SLiM predictions, which can be experimentally validated using proximity labeling strategy owing to their typically low affinity. Finally, this motif knowledge can be used to discover new interactants of your protein of interest using SLiMSearch4. All *in-silico* tools can be accessed via Colab notebooks or webservers: ColabFold: <https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb#scrollTo=kOblAo-xetgx>. AF2\_MiniPAE: <https://colab.research.google.com/drive/1ZAj5gQOP-usd8yAK5uhoLsLH4WPAKxMG>. SLiMsearch4: <https://slim.icr.ac.uk/tools/slimsearch/input>.

In this approach, the predicted SLiM was expressed as a fusion protein (see Methods), and proximity labeling was achieved by fusing the target protein to BioID2, a promiscuous biotin ligase that covalently biotinylates proteins within ~10 nm in living cells [12]. If the SLiM mediates an interaction with the target, the SLiM fusion protein is expected to undergo increased biotinylation, reflecting its proximity to the BioID2-tagged target. Once validated, interacting motifs can be aligned to infer the minimal binding consensus. While identifying the interaction motif of a known partner is informative, the ultimate goal is often to discover new binders. This can be achieved using tools such as SLiMSearch4 [13], which scans the proteomes for additional motif matches. Depending on the specificity of the consensus, this may yield a narrow or broad spectrum of candidates that can be further refined using AF to evaluate whether the motifs adopt a compatible binding interface *in-silico*.

### Illustration of the use of our workflow to find hRSK2-binding SLiMs

To illustrate the practical application of our SLiM discovery workflow, we used it to identify novel motifs binding human RSK2 (hRSK2), a Ser/Thr kinase broadly expressed across tissues and involved in numerous biological processes [14, 15]. Mutations in RSK2 are known to cause Coffin-Lowry syndrome [16] and have been implicated in several other disorders [14], underscoring the

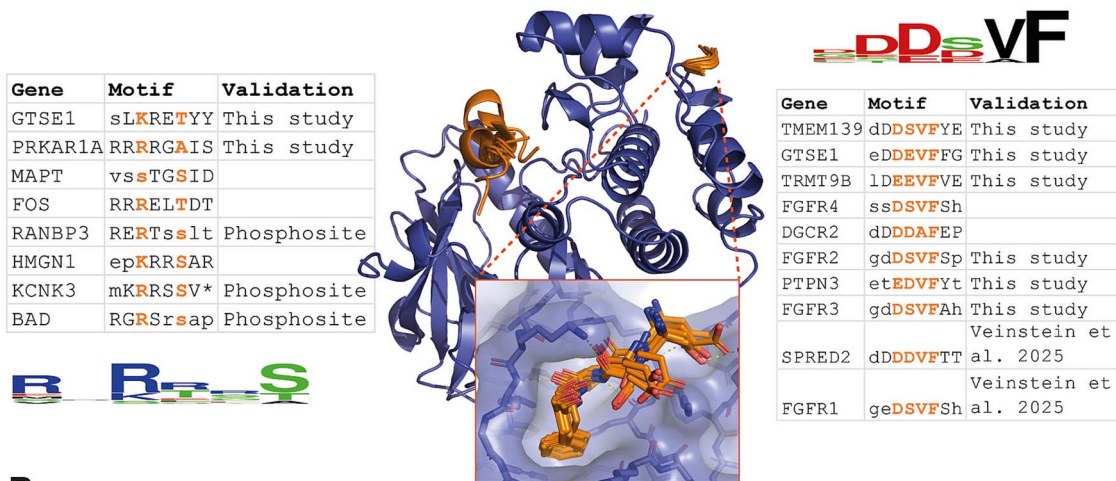
importance of understanding how RSK2 selectively engages its partners through SLiMs.

As outlined in the previous section (Fig. 3), we started by selecting candidate RSK2-binding proteins. For this, we queried the BioGRID and IntAct interaction databases, and, recognizing RSK2's role as a kinase, we also included all known substrates listed for RSK2 in PhosphoSitePlus [17]. This resulted in a list of 220 proteins with less than 1000 amino acids (longer proteins were not considered), which were then individually modeled in binary complexes with RSK2 using AF2. The resulting structures were analyzed using our adapted MiniPAE script, and motifs were filtered using a MiniPAE threshold of 2.85, calculated based on the equation from the previous section, with  $n = 220$ . This analysis yielded 18 predicted SLiMs from 17 proteins.

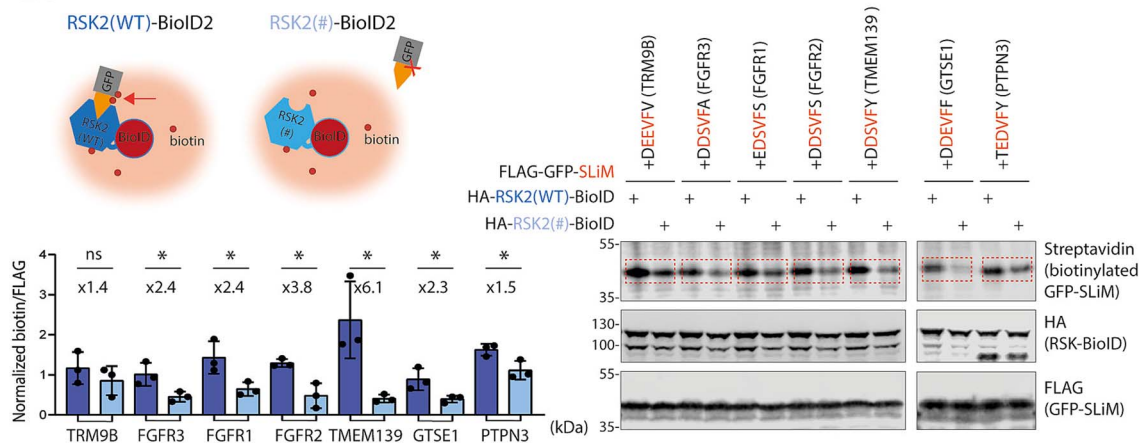
Predicted SLiMs clustered into two distinct binding classes. The first group consisted of 10 'DDVF-like' motifs that engaged a region centered on the KAKLGM loop of RSK2, a surface previously implicated in regulatory interactions and viral hijacking [18, 19]. The second group, comprising 8 'substrate motifs', bound the catalytic cleft of RSK2 and resembled canonical substrate motifs (Fig. 4A, Supplementary data S2).

To experimentally validate these predictions, we selected a subset of representative SLiMs from each class. Among the DDVF-like motifs, two had previously been shown to bind RSK1, consistent with the fact that the closely related RSK1 and RSK2 share the same KAKLGM loop [18, 19]. We focused

A.



B.



C.

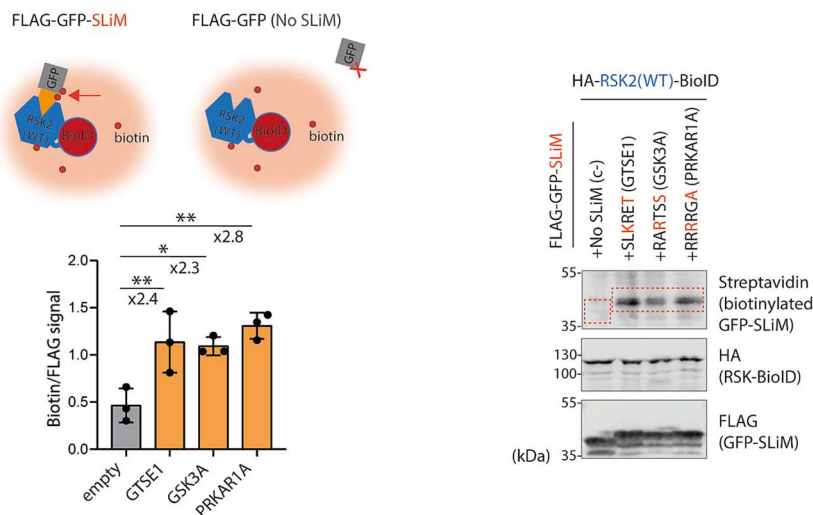


Figure 4. Discovery and validation of hRSK2-binding SLiMs using our AlphaFold-based workflow. (A) AlphaFold2 predictions show two classes of SLiMs binding to hRSK2: Substrate motifs (left) and DDVF-like motifs (right), along with their associated consensus logos. (B–C) Experimental validation of SLiM–RSK2 proximity using a BioID2 proximity labeling. Fold changes and statistical significance are indicated; comparison bars show the groups being tested. Ns = not significant, \* =  $P < 0.05$ , \*\* =  $P < 0.01$ . B: DDVF-like motifs were tested with wild-type or KSEPPY-mutant RSK2. Statistical analysis was performed using a two-tailed unpaired t-test. C: Substrate motifs were tested with wild-type RSK2 and compared to the empty construct. Statistical analysis was performed using one-way analysis of variance (ANOVA). Schematics (top), representative immunoblots (right), and quantifications (left) confirm motif-specific proximity labelling ( $n = 3$ ).

on six new DDVF-like motifs for validation. Given the low affinity of SLiM interactions, we used a proximity labeling strategy involving BioID2 fused to RSK2. Each predicted SLiM was cloned at the C-terminus of a FLAG-GFP construct (see Method), and biotinylation of this protein was compared between cells expressing the wild-type RSK2 or a KAKLGM-to-KSEPPY mutant previously shown to reduce DDVF-mediated interactions. For five out of six SLiMs examined, the KAKLGM-to-KSEPPY mutation in RSK2 significantly reduced biotinylation of the FLAG-GFP-SLiM constructs, confirming RSK2 KAKLGM loop-dependent interactions (Fig. 4B). The degree of biotinylation enrichment by wild-type versus mutant RSK2 varied across SLiMs, ranging from 1.5-fold for PTPN3 to 6.1-fold for TMEM139. Although TRM9B showed a 1.4-fold increase in biotinylation by wild-type RSK2, this difference was not statistically significant at the  $P < 0.05$  level.

For the ‘substrate SLiMs’, five of the eleven predicted motifs were already annotated as RSK2 phosphorylation sites according to PhosphositePlus. We selected two uncharacterized motifs from the remaining six. Validating these was more challenging: substrate motifs are not expected to form stable interactions with kinases due to the transient nature of catalysis, and there are no well-characterized mutations that abolish substrate ‘binding’ without affecting kinase function. Therefore, we assessed proximity by comparing the biotinylation levels of FLAG-GFP-SLiM constructs to those of FLAG-GFP alone. As shown in Fig. 4C, the addition of the motif led to a twofold increased biotinylation, suggesting that our predicted substrate motifs were indeed contributing to the interaction with RSK2. Notably, these motifs matched the known consensus sequence for RSK2 substrates, as curated in PhosphoSitePlus, reinforcing the validity of our predictions.

While these results showed that substrate motifs conferred proximity to RSK2, this alone could not confirm whether these were actually phosphorylated by RSK. To address this, we used an analog-sensitive (AS) kinase system previously developed for RSK1 by Lizcano-Perret *et al.* [20], taking advantage of the high sequence conservation across RSK paralogs [15]. This system enabled the specific detection of phosphorylation events mediated by an analog sensitive RSK (Fig. 5A). For both GTSE1 and GSK3A substrate motifs, phosphorylation of the wild-type SLiM was significantly higher than that of the non-phosphorylatable serine-to-alanine (S > A) mutant, confirming RSK2-dependent phosphorylation (Fig. 5B). Notably, the phosphorylation signal for GTSE1 was markedly weaker than for GSK3A, which could be explained by the fact GSK3A has a canonical substrate motif while GTSE1 does not [15]. We hypothesize that GTSE1’s dual DDVF and substrate motifs may compensate for its non-canonical sequence, facilitating phosphorylation in the context of the full-length protein, an hypothesis supported by the work of Alexa *et al.*, on analogous motifs [21]. In the case of a predicted pseudosubstrate motif in PRKAR1A that naturally contains a non-phosphorylatable alanine at the phospho-acceptor site, we did the converse and introduced an alanine-to-serine mutation and observed phosphorylation of this mutant, further supporting the accuracy of our prediction.

Altogether, our experimental setup allowed to confirm, to various degrees, seven out of eight tested SLiMs, supporting the use of an adaptative MiniPAE threshold for predictive screening. Regarding substrate motifs, our results suggest that AF predictions can extend beyond structural binding to inform on enzymatic modification events, consistent with previous suggestions by Lee *et al.* [5] in the context of protease cleavage site prediction.

## Material and methods

### Benchmark

We downloaded all 2797 interaction instances from the ELM database and retained only those corresponding to *Homo sapiens*–*H. sapiens* interactions. Following the filtering strategy of Bret *et al.* [4], we excluded instances in which the same motif occurred multiple times within a single protein, as this can complicate result interpretations. We also removed a small number of entries for which the reported motif consensus (‘regular expression’) could not be located within the corresponding protein sequence. To avoid redundancy, we kept the shortest instance per ELM class. In addition, to reduce prediction time, we restricted both motif-containing proteins and their partners to sequences shorter than 1000 residues, yielding a final benchmark set of 121 interactions—referred to as the ‘ELM\_dataset’ (see Fig. 1A)—across 161 unique human ELM classes.

Among these, only 26 interactions had no corresponding structural information available in the PDB according to ELM annotations; this subset is referred to as the ‘noPDB\_dataset’.

As a negative control, we generated a ‘nonBinder\_dataset’ by randomly shuffling SLiM–target combinations from the ELM\_dataset, ensuring that each new pair was absent from the original ELM data. To avoid introducing functional motifs by chance, we excluded pairs in which the SLiM matched any regular expression associated with the new target. Finally, we retained only pairs with a total sequence length under 1000 residues and randomly selected 1100 such pairs as likely non-binders. All instances are provided in Supplementary data S1.

### AF prediction

Protein complex structures were predicted using AF2 via ColabFold in multimer V3 mode with default parameters. Predictions using AF3 were performed using the AF Server available at <https://alphafoldserver.com/>.

### MiniPAE and other scoring metrics

#### Regular AF-based metrics

Per-residue pLDDT and PAE matrices were extracted from ColabFold output files. Global confidence metrics (pTM, ipTM) were combined into a MC score:  $MC = 0.2 \cdot pTM + 0.8 \cdot ipTM$  [5]

#### MiniPAE

For binary complex predictions, AF computes a PAE matrix that quantifies the confidence in the relative positioning of different residues. MiniPAE, introduced by Omidi *et al.* [6], represents the lowest (i.e. best) inter-protein PAE value in this matrix. We implemented a simplified MiniPAE script for both AF2 [https://github.com/martino vein/AF2\\_MiniPAE](https://github.com/martino vein/AF2_MiniPAE) and AF3 [https://github.com/martino vein/AF3\\_MiniPAE](https://github.com/martino vein/AF3_MiniPAE) and is available as a pseudocode in Supplementary Data S3.

#### AlphaSLiM score

While pLDDT is commonly employed for SLiM detection, its utility is limited by its increased values in both structured regions and disordered regions upon binding. To overcome this limitation, we developed the AlphaSLiM score, defined as:

$$\text{AlphaSLiM score} = N \cdot (\text{pLDDT}_{\text{bound}} - \text{pLDDT}_{\text{unbound}})$$

where  $N$  is the number interactions calculated by protein-ligand interaction profiler [22]. Motif regions were defined as

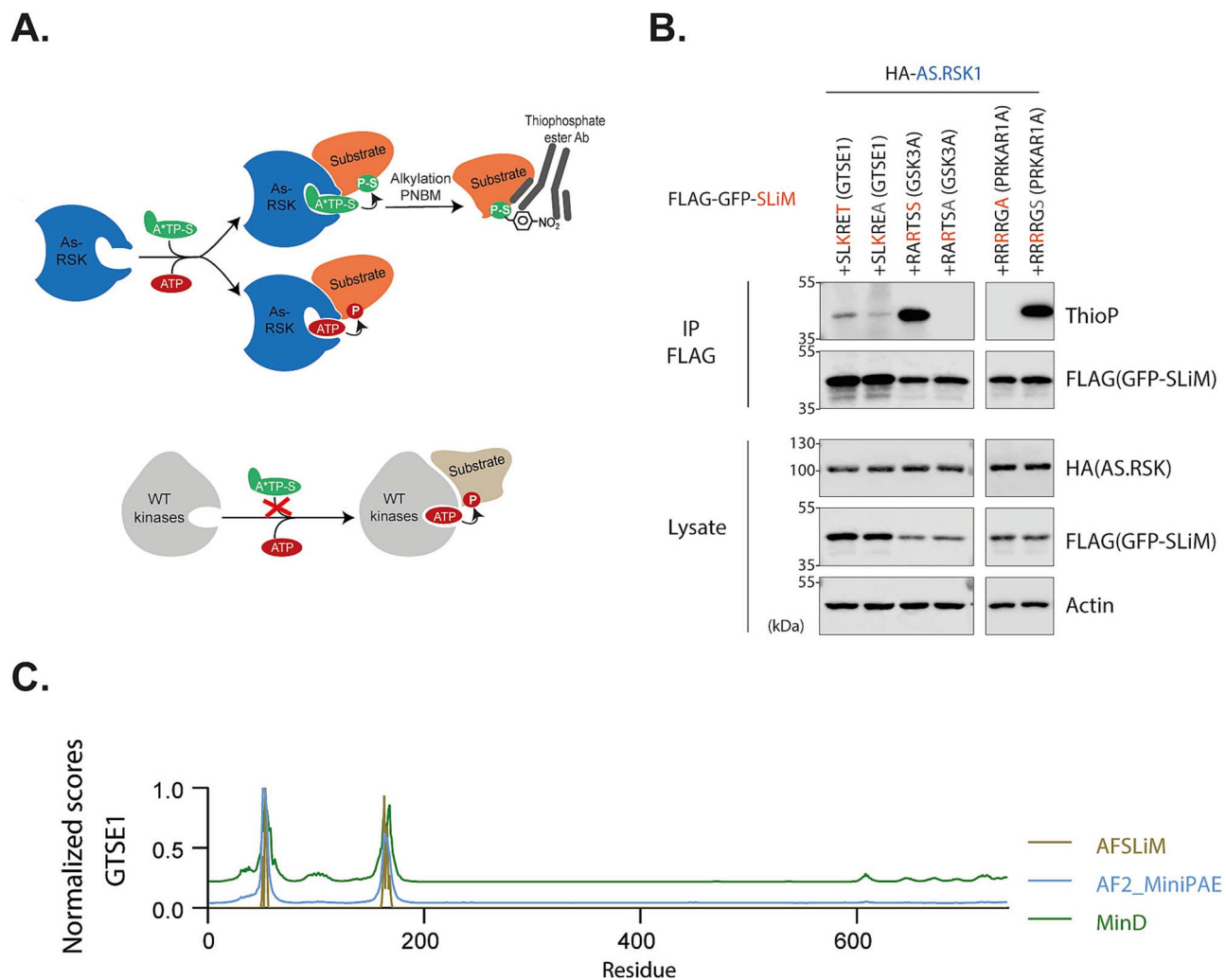


Figure 5. RSK1-dependent phosphorylation of predicted substrate SLiMs using an RSK1 analog-sensitive kinase. (A) Schematic representation of the analog-sensitive kinase system (adapted from Lizcano-Perret et al. [20]). The analog-sensitive (AS) kinase transfers a thiophosphate group from  $N^6$ -bn-ATP- $\gamma$ -S ( $a^*TP-S$ ) to its substrates, enabling subsequent alkylation with PNBM and detection using an anti-thiophosphate ester-specific antibody. (B) Representative immunoblots showing thio-phosphorylation of immunoprecipitated FLAG-GFP-SLiM proteins harboring predicted substrate motifs. Constructs were transfected in HeLa cells expressing HA-tagged AS-*RSK1*. Wild-type motifs exhibit greater RSK1-dependent thiophosphorylation compared to their mutant counterparts ( $n = 2$ ). (C) Different AlphaFold-based scores along GTSE1 sequence when paired with RSK2.

contiguous residues with positive pLDDT differences, scored by their maximum AlphaSLiM score. While MiniPAE offers simpler implementation, AlphaSLiM remains available at <https://github.com/martinovein/AlphaSLiM> and as pseudocode in Supplementary Data S3.

We also used 'MinD', a metric described by Omidi et al. (accessible at <https://github.com/alirezaomidi/AFminD>), and the 'Average\_model score', developed by Schmid et al. available at [9].

Local (per residue) and global (per complex) deltaG were calculated using PDBePISA webserver: <https://www.ebi.ac.uk/pdbe/pisa/>. Buried surface area (BSA) per residue and total interface area were also computed using the same tool.

To facilitate comparative visualization, all scores (AlphaSLiM, MiniPAE, MinD, and BSA) were normalized to a 0–1 scale when indicated. AlphaSLiM and BSA scores were normalized by division with the maximum observed value.

$$\text{Score}_{\text{norm}} = \frac{\text{Score}}{\text{Score}_{\text{max}}}$$

$\text{Score}_{\text{norm}}$  = normalized score

$\text{Score}_{\text{max}}$  = maximum score in the dataset

For MiniPAE and MinD, normalization was performed as follows:

$$\text{Score}_{\text{norm}} = \frac{\text{Score}_{\text{inv}} - \text{Score}_{\text{inv}_{\text{min}}}}{\text{Score}_{\text{inv}_{\text{max}}}}$$

$\text{Score}_{\text{norm}}$  = normalized MiniPAE or MinD score

$\text{Score}_{\text{inv}} = \frac{1}{\text{Score}}$  = inverse of original score

$\text{Score}_{\text{inv}_{\text{min}}}$  = minimum  $\text{Score}_{\text{inv}}$  in the dataset

$\text{Score}_{\text{inv}_{\text{max}}}$  = maximum  $\text{Score}_{\text{inv}}$  in the dataset

### Fitting of power-law model on MiniPAE score

To characterize the relationship between the number of non-binders and the average minimum MiniPAE value, we performed 100 random subsamplings from our nonBinder\_dataset ( $n = 1100$ ). For each iteration, we calculated the minimum MiniPAE value and

then averaged these minima across iterations. We fitted a power-law equation according to:

$$y = a \cdot x^{-b} + c$$

where  $x$  is the number of non-binders,  $y$  is the corresponding average minimum MiniPAE value, and  $a$ ,  $b$ , and  $c$  are model parameters. The fitting was performed in Python using the `curve_fit` function from the `scipy.optimize` module, which applies non-linear least squares optimization.

### Precision-recall curve analysis

Precision-recall (PR) curves and PR-AUC values were computed to assess AlphaFold scores performance on imbalanced data. Positive scores ( $n = 26$ ) and negative scores ( $n = 1100$ ) were extracted from the `noPDB_dataset` and `nonBinder_dataset` respectively. To simulate varying binder-to-non-binder ratios, we randomly sampled 26 (1:1), 130 (1:5), or 650 [1, 25] negative scores and combined them with the 26 positive scores. This process was repeated across 10 independent iterations per ratio, and precision-recall metrics were averaged. Precision and recall were defined as:

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

where TP, FP, and FN are true positives, false positives, and false negatives, respectively. PR curves were generated using `scikit-learn`'s `precision_recall_curve`, and PR-AUC was calculated by trapezoidal integration.

### RSK2 partner selection from BioGRID or IntAct

To identify candidate interaction partners, we searched for the protein of interest in the BioGRID and IntAct databases using its gene name (e.g. `RPS6KA3` for RSK2). The resulting lists of interactors were filtered based on the following inclusion criteria: only human proteins with fewer than 1000 amino acids were retained. Duplicate entries were removed based on protein names and this curated list was then used for downstream analysis.

### Quantification and statistical analysis

ROC curves (Figs 1-2), t-test (Fig. 4B), and one-way ANOVA (Fig. 4C) were calculated as implemented in the Prism 8.0.02 statistical analysis software (GraphPad Software, Inc., San Diego, CA).

### Plasmids, retroviral and lentiviral constructs

Expression plasmids are presented in Supplementary data S4. Note that FLAG tagged proteins expressed by these vectors contain 3xFLAG tag which are referred to as FLAG- in the text and figures. `pBLP10` and `pMD108` plasmids expressing HA-RSK(WT)-BioID2 [23] and HA-RSK(mutant)-BioID2 were kindly provided by Belén Lizcano Perret and Melissa Drappier, respectively. BioID2 [12] coding sequence can be accessed through the Addgene collection (#74223).

All 'FLAG-GFP-SLiM' expression plasmids were created by inserting hybridized oligo between `Bsu36I` and `XbaI` restriction sites within `pMV51` plasmid, a derivative of `pTM952` [24, 25]. SLiMs coding sequences (Supplementary data S4) were generated by Novopro codon optimization tool (<https://www.novoprolabs.com/tools/codon-optimization>).

### Cell culture and transfection

HEK293T [26] and HeLa-BLP32 (expressing analog-sensitive human RSK1 [20]) were maintained in Dulbecco's Modified Eagle Medium (Lonza) supplemented with 10% fetal bovine serum (Sigma), 100 U/ml penicillin and 100  $\mu\text{g/ml}$  streptomycin (Lonza). All cells were cultured at 37 °C in a humidified atmosphere containing 5%  $\text{CO}_2$ .

Cells seeded the day before were transfected using Lipofectamine 2000 (ThermoFisher) according to the manufacturer's instructions. For analog sensitive kinase experiments, cells were incubated with analog sensitive kinase buffer prior to lysis, following the protocol described by Lizcano-Perret et al. [23].

### Biotinylation experiment

Twenty-four hours post-transfection, the HEK293T cells were incubated with 5  $\mu\text{M}$  biotin (diluted in culture medium) for an additional 24 hours at 37 °C. medium was subsequently removed and cells were lysed in Laemmli buffer.

### Immunoprecipitations and western blot analysis

Immunoprecipitation and Western blotting were done as previously described [18]. Briefly, transfected cells were lysed in lysis buffer (Tris-HCl 100 mM pH 8, NaCl 150 mM, NP40 0.5%, EDTA 2 mM, and supplemented with protease/phosphatase inhibitors (Pierce)) and centrifugated at 14,000  $\times g$  for 10 min at 4 °C. Cleared supernatants were incubated with anti-FLAG M2 Magnetic Beads (#M8823, Sigma-Aldrich) with gentle agitation for 4 hours at 4 °C. Magnetic beads were then washed three times with the lysis buffer. Immunoprecipitated proteins were detected by Western-blot analysis antibodies listed in Supplementary data S4.

### Discussion

Although SLiMs are thought to regulate nearly all biological processes, their detection remains notoriously challenging *in vitro* due to the transient and often low-affinity nature of their interactions. *In silico*, early sequence-based motif discovery tools (e.g. MEME, STREME) were considered ill-suited for SLiM prediction from large sequence databases because of the short and degenerate nature of these motifs [27].

However, recent advances in machine learning have changed this landscape. Recent tools such as AIUPred-binding [28] and MoRFchibi 2.0 [29] leverage disorder prediction and deep learning to identify potential SLiM regions with great speed and accuracy (Table 1). Remarkably, MoRFchibi 2.0 has been shown to outperform AF in identifying disordered interaction regions from full-length sequences [29]. Yet, a key limitation is that those approaches do not predict the SLiM's binding partner nor provide structural insight into the interaction, which limits their applicability for screening SLiMs binding a specific target (Table 1).

These approaches are complementary to structure-based SLiM prediction methods: sequence-based tools enable rapid, large-scale scanning of disordered regions, whereas structure-based approaches like AF can pinpoint specific binding partners and generate structural hypotheses for SLiM-target interaction interfaces.

Our study extends previous evaluations of AF2 for SLiM recovery by including AlphaFold3 and leveraging a structure-independent, ELM-derived benchmark of 26 interactions entirely absent from PDB homology. This benchmark demonstrated that MiniPAE, computed on AF2 predictions, performs comparably to both MinD and our AlphaSLiM scoring function. While all

Table 1. Comparison of different SLiM discovery methods.

	Method	Examples	Time	Cost	'Receptor' proteins to screen	'SLiM' proteins screened	Structural insight
In silico	Neural network sequence-based predictors	AIUPred-binding [27] or MorfChibi 2.0 (28)	Seconds	Free	na	Proteome-wide	no
In vitro	AlphaFold	Colabfold	Days	Free/+	1–10	1–1000	yes
	High throughput Phage display	ProPD [1]	Months	++	1–100	Proteome-wide	no
	Proximity labelling	BioID2	Weeks	+	1	1–100	no

three are effective, MiniPAE's simpler implementation makes it particularly practical for general users, avoiding the large extra outputs of MinD and the need for additional monomer predictions required by AlphaSLiM.

Despite promising results, benchmarking revealed key limitations. In unbalanced datasets mimicking real-world screening—where non-binders vastly outnumber binders—precision decreases exponentially. This finding precludes blind SLiM screening with AlphaFold and instead suggests starting from interaction datasets already enriched for true-binders like those available on BioGRID and Intact. To mitigate false positives, we proposed adaptive score thresholds tailored to specific screening strategies. Alternatively, AF predictions could be combined with post-hoc rescoring tools such as SPOC [9], though current implementations are limited to human reference proteins.

Finally, users planning SLiM screening experiments should account for AF's intrinsic constraints. The model performs less reliably for certain protein classes (e.g. coiled-coils, antibody-antigen complexes, and some viral proteins), and prediction time increases exponentially with sequence length.

MiniPAE scores derived from AF3 predictions were slightly lower than those from AF2, echoing recent observations in peptide-protein interaction benchmarking [30]. The transition from AF2's transformer-based architecture to AF3's diffusion-based generative model [7] may underlie this difference. These findings support AF2 as the more practical choice for this application, especially given current throughput limitations on the AF3 server.

From a biological perspective, the generalizability of this workflow across SLiM types appears promising. The workflow successfully identified two RSK2-binding SLiM classes in this work with distinct interaction properties—one mediating transient, substrate-like binding and the other forming more stable complexes—indicating its capability to detect diverse SLiM types. We identified 13 previously undescribed RSK2-binding SLiMs, and using our sensitive proximity labelling system, we confirmed 8/9 tested predictions experimentally. Among these, we found that GSTE1 interacts with RSK2 via two distinct SLiMs: a substrate-like motif divergent from the canonical phosphorylation site, and a DDVF motif, which likely engages a different region near the kinase active site. We hypothesized that this dual SLiM strategy may facilitate the phosphorylation of non-canonical substrates, expanding the functional repertoire of the kinase [21]. However, the functional relevance of these motifs remains to be confirmed in the context of full-length GSTE1, as our current evidence is limited to peptide-based assays.

High-throughput phage display screening of SLiMs has previously achieved a recall of 23% [1], but its elevated cost and substantial protein purification workload limit broader adoption (Table 1). In contrast, our workflow offers a comparable recall of 30% with minimal infrastructure requirements (Table 1). The rise

of ColabFold has democratized the use of structure prediction, allowing any researcher to perform interaction predictions without the need for local installation. Building on this, we provide a Google Colab notebook for SLiM identification across multiple AF predictions—eliminating installation barriers entirely. Together, these tools aim to make AF-based SLiM prediction accessible even to researchers without coding expertise.

#### Key Points

- Various AlphaFold2/3 scoring metrics were systematically benchmarked for short linear motif (SLiM) detection using a unique dataset of 26 interactions with no PDB homology, expanding on previous AlphaFold2 evaluations.
- Given our experimental background, we highlight a critical and often-overlooked limitation, showing that AlphaFold lacks specificity when screening for rare SLiMs.
- A user-friendly and cost-effective AlphaFold-based workflow was specifically designed to identify novel SLiMs for a protein of interest.
- This computational workflow was paired with a sensitive *in-vitro* assay to offer a complete solution for low-affinity SLiM discovery and validation, allowing the identification of 13 previously uncharacterized SLiMs interacting with RSK kinases.
- Together, our workflow and associated validation assay offer an integrated pipeline for the discovery and validation of SLiM-mediated protein-protein interactions.

## Acknowledgements

We thank Alireza Omid for the original miniPAE script [6], Belen Lizcano-Perret for BLP32 cells, the BLP10 plasmid, and Fig 5A's kinase illustration [20, 23], Melissa Drappier for the MD108 plasmid, Laurent Gatto for benchmarking guidance, and Nicolas Papadopoulos for advice on AlphaFold screenings. We used ChatGPT to support code writing and language editing.

## Author contributions

M.V., F.S., and T.M. (Conceptualization), M.V., B.I., and V.J. (Methodology), R.H. and B.I. (Software), M.V. (Formal analysis), M.V. and F.S. (Writing), F.S. and T.M. (Supervision and Funding acquisition).

## Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Conflict of interest: All authors declare that they have no conflicts of interest.

## Funding

MV is the recipient of an Aspirant fellowship from the Fonds de la recherche scientifique-FNRS. F.S. is supported by an INRS starter grant. Work in TM's lab was supported by FNRS (PDR T.0154.23), EOS (EOS ID: 30981113 and 40007527), and Loterie Nationale through support to the de Duve Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Data availability

All data are available for download and biological material is freely available upon request. All benchmark AlphaFold2 data (ELM\_dataset including noPDB\_dataset, ~200Gb in 5 parts) is available at: <https://zenodo.org/uploads/15260141>, <https://zenodo.org/uploads/15260224>, <https://zenodo.org/uploads/15273252>, <https://zenodo.org/uploads/15273254>, <https://zenodo.org/uploads/15273263>.

The monomeric predictions additionally required for AlphaSLiM are available at: <https://zenodo.org/uploads/15236595>.

AlphaFold3 predictions for noPDB\_dataset are available at: <https://zenodo.org/uploads/15236655>.

hRSK2 SLiM predictions are available at <https://zenodo.org/uploads/15235709>. The monomeric predictions required for AlphaSLiM are available at:

<https://doi.org/10.5281/zenodo.15236154>

## References

1. Benz C, Ali M, Krystkowiak I. *et al.* Proteome-scale mapping of binding sites in the unstructured regions of the human proteome. *Mol Syst Biol* 2022;**18**:e10584. <https://doi.org/10.15252/msb.202110584>
2. Kumar M, Michael S, Alvarado-Valverde J. *et al.* ELM—the eukaryotic linear motif resource—2024 update. *Nucleic Acids Res* 2024;**52**:D442–55. <https://doi.org/10.1093/nar/gkad1058>
3. Zeke A, Alexa A, Reményi A. Discovery and characterization of linear motif mediated protein-protein complexes. *Adv Exp Med Biol* 2024;**3234**:59–71. [https://doi.org/10.1007/978-3-031-52193-5\\_5](https://doi.org/10.1007/978-3-031-52193-5_5)
4. Bret H, Gao J, Zea DJ. *et al.* From interaction networks to interfaces, scanning intrinsically disordered regions using AlphaFold2. *Nat Commun* 2024;**15**:597. <https://doi.org/10.1038/s41467-023-44288-7>
5. Lee CY, Hubrich D, Varga JK. *et al.* Systematic discovery of protein interaction interfaces using AlphaFold and experimental validation. *Mol Syst Biol* 2024;**20**:75–97. <https://doi.org/10.1038/s44320-023-00005-6>
6. Omidi A, Møller MH, Malhis N. *et al.* AlphaFold-multimer accurately captures interactions and dynamics of intrinsically disordered protein regions. *Proc Natl Acad Sci U S A* 2024;**121**:e2406407121. <https://doi.org/10.1073/pnas.2406407121>
7. Abramson J, Adler J, Dunger J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;**630**:493–500. <https://doi.org/10.1038/s41586-024-07487-w>
8. Orand T, Jensen MR. Binding mechanisms of intrinsically disordered proteins: insights from experimental studies and structural predictions. *Curr Opin Struct Biol* 2025;**90**:102958. <https://doi.org/10.1016/j.sbi.2024.102958>
9. Schmid EW, Walter JC. Predictomes, a classifier-curated database of AlphaFold-modeled protein-protein interactions. *Mol Cell* 2025;**85**:1216–1232.e5. <https://doi.org/10.1016/j.molcel.2025.01.034>
10. Oughtred R, Rust J, Chang C. *et al.* The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci* 2021;**30**:187–200. <https://doi.org/10.1002/pro.3978>
11. del Toro N, Shrivastava A, Ragueneau E. *et al.* The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res* 2021;**50**:D648–53. <https://doi.org/10.1093/nar/gkab1006>
12. Kim DI, Jensen SC, Noble KA. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell* 2016;**27**:1188–96. <https://doi.org/10.1091/mbc.E15-12-0844>
13. Krystkowiak I, Davey NE. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. *Nucleic Acids Res* 2017;**45**:W464–9. <https://doi.org/10.1093/nar/gkx238>
14. Anjum R, Blenis J. The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* 2008;**9**:747–58. <https://doi.org/10.1038/nrm2509>
15. Romeo Y, Zhang X, Roux PP. Regulation and function of the RSK family of protein kinases. *Biochem J* 2012;**441**:553–69. <https://doi.org/10.1042/BJ20110289>
16. Delaunoy JP, Dubos A, Marques Pereira P. *et al.* Identification of novel mutations in the RSK2 gene (RPS6KA3) in patients with Coffin-Lowry syndrome. *Clin Genet* 2006;**70**:161–6. <https://doi.org/10.1111/j.1399-0004.2006.00660.x>
17. Hornbeck PV, Zhang B, Murray B. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 2015;**43**:D512–20. <https://doi.org/10.1093/nar/gku1267>
18. Veinstein M, Stroobant V, Wavreil F. *et al.* The 'DDVF' motif used by viral and bacterial proteins to hijack RSK kinases mimics a short linear motif (SLiM) found in proteins related to the RAS-ERK MAP kinase pathway. *PLoS Pathog* 2025;**21**:e1013016. <https://doi.org/10.1371/journal.ppat.1013016>
19. Sorgeloos F, Peeters M, Hayashi Y. *et al.* A case of convergent evolution: several viral and bacterial pathogens hijack RSK kinases through a common linear motif. *Proc Natl Acad Sci U S A* 2022;**119**:e2114647119. <https://doi.org/10.1073/pnas.2114647119>
20. Lizcano-Perret B, Vertommen D, Herinckx G. *et al.* Identification of RSK substrates using an analog-sensitive kinase approach. *J Biol Chem* 2024;**300** (3):105739. <https://doi.org/10.1016/j.jbc.2024.105739>
21. Alexa A, Sok P, Gross F. *et al.* A non-catalytic herpesviral protein reconfigures ERK-RSK signaling by targeting kinase docking systems in the host. *Nat Commun* 2022;**13**:472. <https://doi.org/10.1038/s41467-022-28109-x>
22. Adasme MF, Linnemann KL, Bolz SN. *et al.* PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Res* 2021;**49**:W530–4. <https://doi.org/10.1093/nar/gkab294>
23. Lizcano-Perret B, Lardinois C, Wavreil F. *et al.* Coronavirus leader proteins retarget RSK kinases toward alternative substrates to perturb nucleocytoplasmic traffic. *PLoS Pathog* 2022;**18**:e1011042. <https://doi.org/10.1371/journal.ppat.1011042>
24. Cesaro T, Hayashi Y, Borghese F. *et al.* PKR activity modulation by phosphomimetic mutations of serine residues located three aminoacids upstream of double-stranded RNA binding motifs. *Sci Rep* 2021;**11**:9188. <https://doi.org/10.1038/s41598-021-88610-z>

25. Follenzi A, Ailles LE, Bakovic S. et al. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat Genet* 2000;**25**:217–22. <https://doi.org/10.1038/76095>
26. DuBridge RB, Tang P, Hsia HC. et al. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol Cell Biol* 1987;**7**:379–87.
27. Haslam NJ, Shields DC. Profile-based short linear protein motif discovery. *BMC Bioinformatics* 2012;**13**:104. <https://doi.org/10.1186/1471-2105-13-104>
28. Erdős G, Deutsch N, Dosztányi Z. AIUPred – binding: energy embedding to identify disordered binding regions. *J Mol Biol* 2025;**437**:169071. <https://doi.org/10.1016/j.jmb.2025.169071>
29. Malhis N, Jacobson M, Gsponer J. MoRFchibi SYSTEM: software tools for the identification of MoRFs in protein sequences. *Nucleic Acids Res* 2016;**44**:W488–93. <https://doi.org/10.1093/nar/gkw409>
30. Manshour N, Ren JZ, Esmaili F. et al. Comprehensive evaluation of AlphaFold-multimer, AlphaFold3 and ColabFold, and scoring functions in predicting protein-peptide complex structures. *bioRxiv [Internet]* 2024. Available from: <https://www.biorxiv.org/content/early/2024/11/11/2024.11.11.622992>