

SOFTWARE

Open Access



SPServer: split-statistical potentials for the analysis of protein structures and protein–protein interactions

Joaquim Aguirre-Plans¹, Alberto Meseguer¹, Ruben Molina-Fernandez¹, Manuel Alejandro Marín-López¹, Gaurav Jumde¹, Kevin Casanova¹, Jaume Bonet², Oriol Fornes³, Narcis Fernandez-Fuentes^{4,5} and Baldo Oliva^{1*} 

*Correspondence:

baldo.oliva@upf.edu

¹ Structural Bioinformatics Lab, Department of Experimental and Health Science, Universitat Pompeu Fabra, 08003 Barcelona, Catalonia, Spain
Full list of author information is available at the end of the article

Abstract

Background: Statistical potentials, also named knowledge-based potentials, are scoring functions derived from empirical data that can be used to evaluate the quality of protein folds and protein–protein interaction (PPI) structures. In previous works we decomposed the statistical potentials in different terms, named Split-Statistical Potentials, accounting for the type of amino acid pairs, their hydrophobicity, solvent accessibility and type of secondary structure. These potentials have been successfully used to identify near-native structures in protein structure prediction, rank protein docking poses, and predict PPI binding affinities.

Results: Here, we present the SPServer, a web server that applies the Split-Statistical Potentials to analyze protein folds and protein interfaces. SPServer provides global scores as well as residue/residue-pair profiles presented as score plots and maps. This level of detail allows users to: (1) identify potentially problematic regions on protein structures; (2) identify disrupting amino acid pairs in protein interfaces; and (3) compare and analyze the quality of tertiary and quaternary structural models.

Conclusions: While there are many web servers that provide scoring functions to assess the quality of either protein folds or PPI structures, SPServer integrates both aspects in a unique easy-to-use web server. Moreover, the server permits to locally assess the quality of the structures and interfaces at a residue level and provides tools to compare the local assessment between structures.

Server address: <https://sbi.upf.edu/spserver/>.

Keywords: Protein structure evaluation, Protein structure quality assessment, Protein structure prediction, Protein–protein interaction, Protein–protein evaluation, Knowledge-based potential

Background

Three-dimensional (3D) structures of proteins and protein–protein interactions (PPIs) are essential to understand most biochemical functions of cells and living organisms. Yet, the amount of experimentally determined 3D structures is limited, especially for protein complexes. Structural models derived by computational methods can be used



© The Author(s) 2020. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

to close the gap between the number of sequences and structures. In the recent CASP13 competition, we have observed a dramatic progress in the quality of the template-free models made by novel computational methods involving deep learning techniques [1]. However, these methods need to be complemented by evaluation methods to know the margins of accuracy when we study the role of structural models in a biological system [2].

Evaluation methods can be classified into two categories: single- and multiple-model methods. Single-model methods only require one model as input, whereas multiple-model methods require several. The latter ones take advantage of the similarity between the distinct models to evaluate them, but they are not based on the properties of the model itself. In contrast, single-model methods are often based on the geometric and energetic analysis of the model coordinates, although some of them may also use additional information (e.g. for evolutionary related proteins) [3, 4].

For single-model methods, the most common approach is to use knowledge-based potentials, i.e. scoring functions derived from the analysis of empirical data [5]. Several computational methods have been implemented from knowledge-based potentials [6–8]. Split-Statistical Potentials (SPs) are knowledge-based potentials that consider the frequency of pairs of residues in contact and include their structural environment, such as solvent accessibility and type of secondary structure. Previously, we demonstrated that SPs can be used to: (1) identify near-native protein decoys in structure prediction [9]; and (2) rank protein–protein docking poses [10, 11]. SPs compared favorably against 115 scoring functions on a docking decoy benchmark [12] and were successful at predicting binding energies of PPIs without requiring the native structures of the complexes [13].

Many scoring methods have been proposed to assess the quality of protein fold models [6–8, 14–18]. However, very few can be easily accessed as web servers by the non-specialized user. In most cases, the web servers have a reduced input flexibility (i.e. only accept models in PDB format, require chain identifiers and protein sequences, or do not accept multiple structures) and a complicated visualization of the results (i.e. do not permit to download results or do not have 3D visualization capabilities).

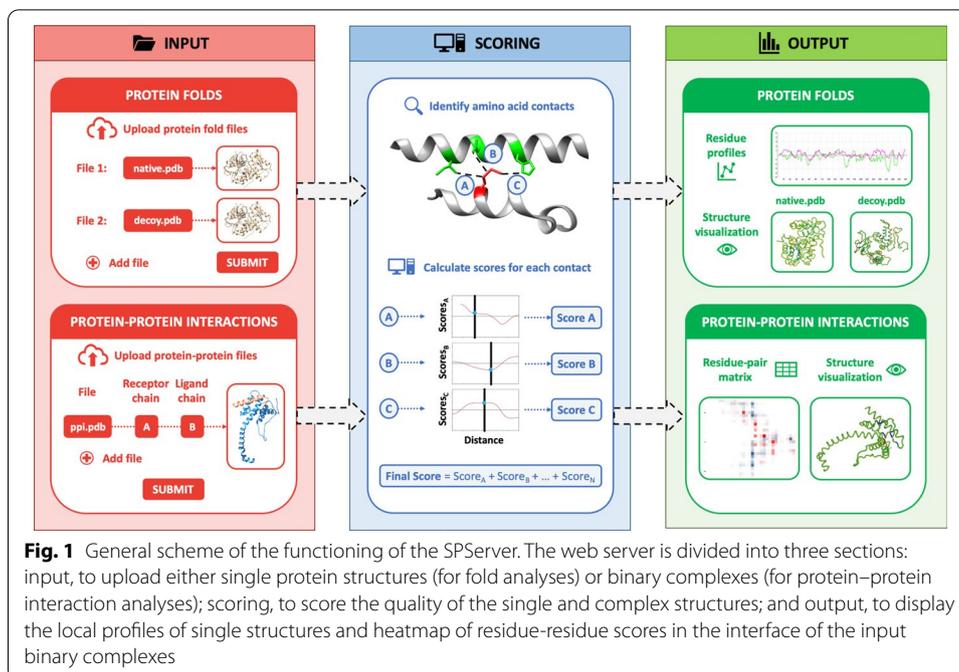
Here, we present the Split-Statistical Potentials Server (SPServer) featuring our SPs for the evaluation of protein structures and PPIs. The web server has been designed to facilitate its use and the interpretation of results. When analyzing protein folds, the server returns global scores and shows score profiles along the protein sequence to identify potentially problematic regions in the structure. When analyzing PPIs, the server returns global scores and score maps of the interfaces. The SPServer identifies stabilizing and disrupting residue pairs that can be used as starting point for follow up protein engineering.

Implementation

The overall implementation of the web server is summarized in Fig. 1 and explained in detail as follows:

Input

As input, users have to provide the structures of one or more proteins or protein complexes. The server input is flexible; users can provide either PDB structures, mmCIF



files or compressed directories containing the structures to analyze. Users also have to select the parameter used to define residue contacts (i.e. 12 Å cut-off between their β -carbons—option C β —or 5 Å between any atom of each residue—option MIN—). Often the structures used as input are produced by modelling or fold prediction approaches, because we are interested in checking the quality of models rather than the quality of experimental structures. In the case of structures of single proteins or folds, the most common methods to produce them are by homology modelling (e.g. by MODELLER [19]), remote homology (e.g. by PHYRE [20] or FUGUE [21]), by threading and ab initio fold prediction (e.g. by I-TASSER [22], THREADER [23], or in particular for sequences in CASP13 using AlphaFold [24]), or protein structure design (e.g. with Rosetta [25]). For protein–protein interactions the structures may be produced by template homology (e.g. from Interactome3D [26], PrePPI [27] or MODPIN [28]), template docking (e.g. by ICM [29]), docking (e.g. by pyDOCK [30], FTDOCK [31], V-D2OCK [32], PatchDock [33] or ZDOCK [34]) or directed docking (e.g. RosettaDock [25] and HADDOCK [35]).

Scoring

The first step of the scoring process is to identify the contacts between residues from the same protein (to score protein folds) or from different proteins (to score PPIs). These contacts consider the amino acids type, the distance between them, and environmental features such as the type of secondary structure or the degree of exposure of the amino acids. SPs provide a score for each one of these contacts. We obtain the score of a structure by performing the sum of scores of all its contacts. We can also get the scores of individual amino acids by performing the sum of scores of all the contacts of that residue. This can be used to define a score profile along the protein

sequence. Residue scores can be averaged using a sliding window of size defined by the user along the protein sequence in order to smooth the profile.

We defined SPs in previous works [9, 10] using the description of a potential of mean force (PMF), say the features describing an amino acid are defined by θ , with: $\theta = (\text{secondary structure, polar character, degree of exposure})$. Then we define the potentials as in Eqs. 1–5:

$$PMF_{PAIR}(a, b) = -k_B T \log \left(\frac{P(a, b|d_{ab})}{P(a)P(b)P(d_{ab})} \right) \quad (1)$$

$$PMF_{LOCAL}(a, b) = k_B T \log \left(\frac{P(a|\theta_a)P(\theta_a)}{P(a)} \right) + k_B T \log \left(\frac{P(b|\theta_b)P(\theta_b)}{P(b)} \right) \quad (2)$$

$$PMF_{3D}(a, b) = k_B T \log(P(d_{ab})) \quad (3)$$

$$PMF_{3DC}(a, b) = k_B T \log \left(\frac{P(\theta_a, \theta_b|d_{ab})}{P(\theta_a, \theta_b)} \right) \quad (4)$$

$$PMF_{S3DC}(a, b) = -k_B T \log \left(\frac{P(a, b|d_{ab}, \theta_a, \theta_b)P(\theta_a, \theta_b)}{P(a, b|\theta_a, \theta_b)P(\theta_a, \theta_b|d_{ab})} \right) \quad (5)$$

with k_B the Boltzmann constant, T the standard temperature (300 K), θ_a , and θ_b the features of amino acids a and b , and d_{ab} the distance between both residues. The terms $P(\cdot)$ denote the probabilities of observing interacting pairs (with or without conditions). For instance, $P(a, b|d_{ab})$ is the conditional probability that residues a, b interact at distance smaller than or equal to d_{ab} , and $P(d_{ab})$ is the probability of finding any pair of residues interacting at distance smaller than or equal to d_{ab} .

The scores *PAIR*, *ELOCAL*, *E3D*, *E3DC*, and *ES3DC* are obtained by summing the PMF with the corresponding subindex of each pair of interacting residues a, b , either of the same protein (for fold) or between two interacting proteins (for PPIs), as in Eq. 6:

$$E = \sum_{a, b} PMF(a, b) \quad (6)$$

We proved [9] that the classical statistic potential, *PAIR*, can be approximated to:

$$PAIR = ES3DC - E3DC + E3D - ELOCAL + \varepsilon \quad (7)$$

With a residual ε that accounts for the reference state and becomes noise centered at 0 upon normalization (i.e. when transformed in Z -scores, see further). Hence, given that *E3D* nullifies when normalizing the scores and ε is irrelevant, we define another score, *ECOMB*, as:

$$ECOMB = ES3DC - E3DC - ELOCAL \quad (8)$$

Furthermore, these potentials can be used to generate a profile per amino acid position along the sequence by summing the energies of the contacts of each residue.

In conclusion, the SPServer has 6 types of SPs available that differ on the environmental features considered for the contact definition: (1) *ES3DC* considers residue frequencies along distances and their environments (i.e. hydrophobicity of each amino acid, solvent accessibility and secondary structure); (2) *E3DC* considers frequencies along distances of pairs referred by the hydrophobicity of the amino acids and the rest of their environments; (3) *PAIR* considers amino acid frequencies along distances; (4) *ELOCAL* considers amino acid frequencies on a particular environment; (5) *E3D* considers the frequencies of any pair of residues along distances; and finally, (6) *ECOMB* combines *ES3DC*, *ELOCAL* and *E3DC* scores [9]. Additionally, Z-scores are provided for each one of these scoring functions by normalizing the scores with respect to the average and standard deviation of 1000 random sequences with the same structure. Similarly, scoring profiles can also be transformed into Z-scoring profiles by normalizing with respect to the 20 possible amino acids in each position. As calculated, scores are proxy measures for energy, and thus, the lowest the score is, the closer the model is to the native-like structure.

Output for protein folds

For a set of protein folds, the SPServer outputs: (1) the global scores (raw and normalized) of all SPs; and (2) the scoring profile per residue (local scores) along the protein sequence. Global scores account for the overall quality of structural models, while per-residue score plots pinpoint problematic regions of the models that likely have either a wrong conformation or contacts with a wrongly modelled region.

Output for protein–protein interactions

For PPIs, the server outputs: (1) global scores for the quality of the interface between the two interacting proteins; (2) a measure of the penetration between two proteins to assess for steric clashes at the interface; and (3) interface maps with the scores of residue contact-pairs between the two proteins. Global scores inform on the overall quality of the interaction (i.e. for ranking docking poses). The measurement of steric hindrances is indicated in a color legend depending on the relevance of the clashes (see Additional file 1: Data and Additional file 2: Figure S1 for details). Finally, interface maps allow for detailed exploration of the protein interfaces at residue level. The server also provides different tools to smooth and compare interface maps.

Results and discussion

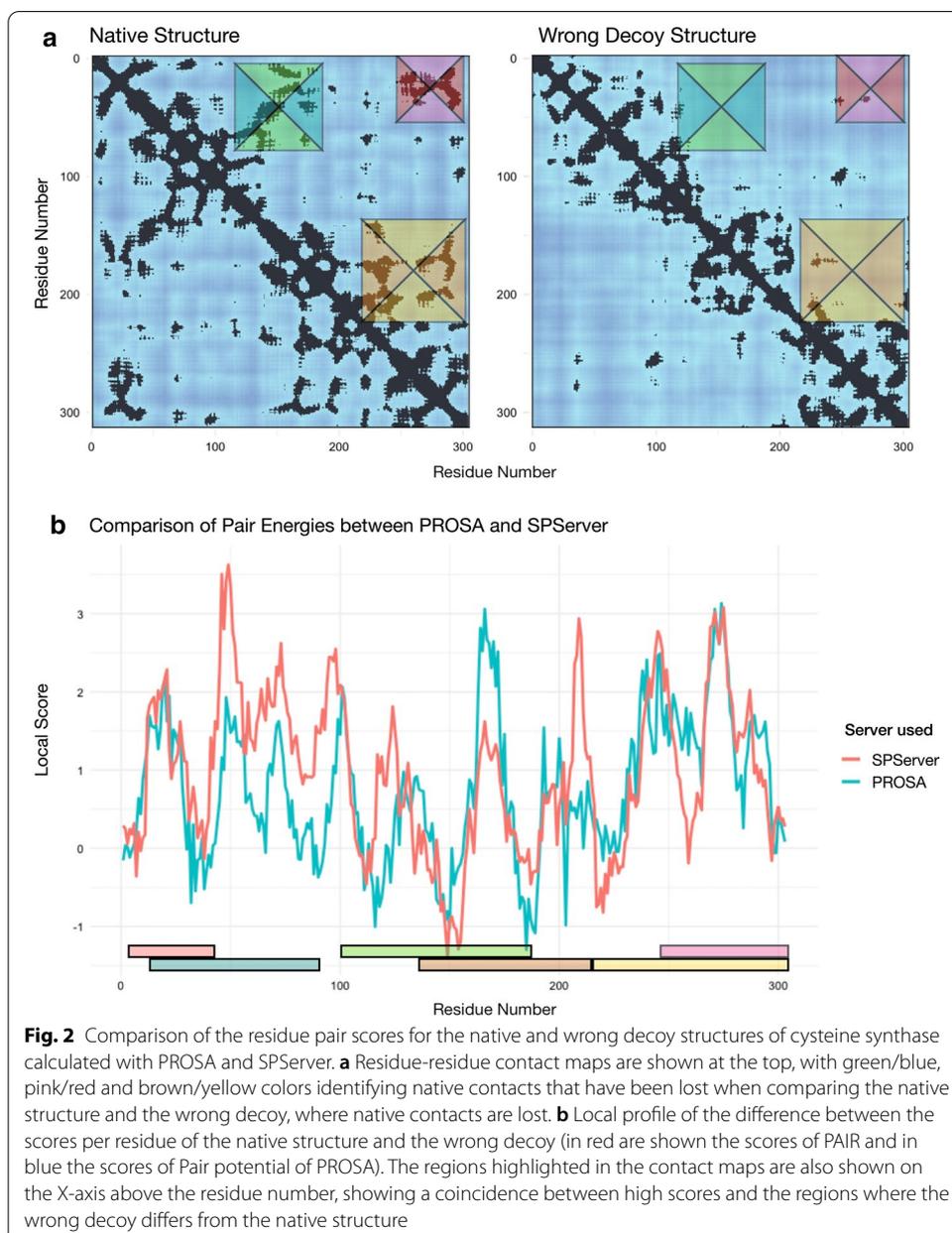
Case study 1: Evaluation of the structural models of Cysteine synthase A

We compared the native structure of Cysteine synthase A from *E. coli* with two decoys of predicted structures: a near-native structure and a wrong decoy. All structures were retrieved from the CASP12 dataset (codes T0861, T0861TS275_2 and T0861TS321_1) [36]. The global scores rank the native structure with the lowest score, followed by the near-native and the wrong decoy (see Additional file 13: Table S1). Local score profiles of the native and the near-native structures are very similar, while the profile of the wrong decoy is different (see Additional file 3: Figure S2 and Additional file 4: Figure S3). Moreover, we compared the results of SPServer PAIR potential with a standard statistical potential (PROSA [6]). Both potentials show similar differences between the profiles

of the native structure and the wrong decoy (Pearson correlation coefficient = 0.50), and highlight the residue-residue contact areas corresponding with wrongly modelled regions of the decoy structure (see Fig. 2).

Case study 2: Mutation in the interaction between BAX and BID

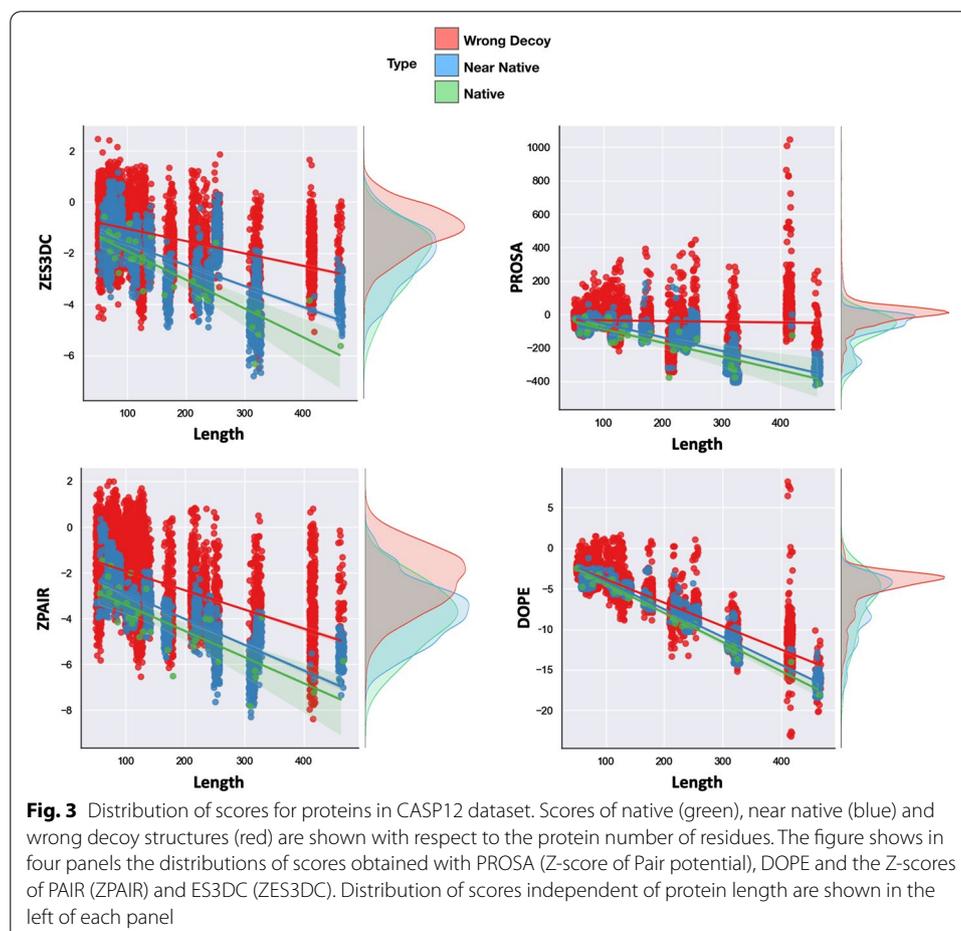
The interaction of BAX with BID mediates the insertion of BAX in the outer mitochondrial membrane, which induces apoptosis [37]. The BAX variant G108V has been associated with Burkitt Lymphoma [38]. We analyzed the interaction BAX-BID in its native form and the G108V variant (mutant form) generated with Modeller [19]. At a global level, only two of SPs are slightly higher for the mutant (i.e. PAIR, ES3DC and their



respective Z-scores) while the rest remain unaffected (see Additional file 14: Table S2). However, the analysis of the interface identifies the detrimental effect of the mutation, as observed in the region around residues 108–110 of BAX (see Additional file 5: Figure S4).

Evaluation of the SPServer global and residue scores on the CASP12 benchmark

We test the SPs of the SPServer on the CASP12 [36] benchmark curated by López-Blanco et al. [39] (Additional file 17: Table S5). We classify the decoys of the benchmark as near-native ($GDT_TS \geq 65\%$, as defined in [40]) and wrong ($GDT_TS < 65\%$). The final CASP12 benchmark contains 9,977 structures, of which 2,100 were classified as near-native and 7,845 as wrongly modelled, and 32 were the native structure. We compare SPServer global and local scores with those from two standard scoring programs: PROSA [6] and DOPE [41]. In Fig. 3, we show the distributions of different scores for wrongly modelled decoys, near-native decoys and native structures in the CASP12 benchmark for proteins with different length. The scoring functions distinguish between native and non-native structures, assigning lower scores to native, higher scores to near-native and much higher to wrong decoy conformations. For proteins longer than 200 residues, all scoring approaches clearly separate native, near-native and wrong



conformations. However, the scores of PROSA (Z-score of Pair potential), ZES3DC (Z-score normalized ES3DC) and ZPAIR (Z-score of PAIR) are optimal to distinguish between native and non-native structures.

In the Additional file 1: Data, we include the pairwise correlations between the global (full protein) and local (per residue) scores of the SPSEver scoring functions PAIR and ES3DC, and the scores of PROSA (Pair potential) and DOPE. The Pearson correlation coefficients between the potentials ZPAIR and ZES3DC and the state-of-the-art potentials PROSA and DOPE are higher than 0.6 (ranging between 0.6 and 0.72, see Additional file 15: Table S3 and Additional file 6: Figure S5). We also compared the local scores (profiles per residue) of the different scoring functions. The SPSEver profiles with score PAIR are correlated with the profiles using DOPE (0.57) and PROSA (0.38) (see Additional file 16: Table S4 and Additional file 12: Figure S11).

Additionally, we compare the global Z-scores of SPSEver with three quality metrics used as reference in CASP: Template Modelling (TM) score [42], Global Distance Test (GDT_TS) [43] and Quality Control Score (QCS) [44]. TM score and GDT_TS measure the quality of a model based on its similarity with the native structure. In contrast, QCS measures the quality of the model based on structural features such as the position of its secondary structure elements. Additional file 15: Table S3 and Additional file 6: Figure S5 compare ZPAIR and ZES3DC global scores with TM, GDT_TS and QCS (Pearson correlations range between -0.44 and -0.58). Our scores compete with other scores, such as the Z-score of PROSA or the global score of DOPE, showing similar Pearson correlations with both (ranging between -0.1 and -0.47), proving their utility to detect the right fold among several decoys. The comparison of scores for all the CASP12 structures can be easily visualized as scatter plots in Additional files 7–11: Figures S6–S10.

Comparison of the SPSEver interface with other protein scoring web servers

We compared the SPSEver in terms of input flexibility, user-friendliness, speed and intuitive visualization of results with other state-of-the-art functional web servers for protein fold assessment (ANOLEA [14], MODFOLD6 [18], ProQ3D [17], ProSA-web [6], QMEAN [16], Verify 3D [15], VoromQA [8]). SPSEver, ANOLEA [14], PROSA-web [6] and QMEAN [16] use statistical potentials. ModFOLD6 [18] and ProQ3D [17] combine several structural features and outputs from 3rd party software into neural networks. QMEAN [16] and VERIFY 3D [15] analyze local structural features such as the secondary structure, the degree of exposure and the degree of polarity for each amino acid. VoromQA [8] analyzes contact regions based on the study of van der Waals radius through Voronoi tessellations. The comparison is summarized in Table 1.

In terms of input flexibility, the SPSEver accepts both PDB and mmCIF formats, inputs with single or multiple structures, and does not require the sequence or the identifiers of the protein chains because it handles everything automatically. In contrast, only ProQ3D and QMEAN accept mmCIF format, and only MODFOLD6, ProQ3D, QMEAN and VoromQA accept multiple structures.

In terms of scoring calculation, all the web servers offer both global and local scores in short time. The only web server requiring some extra time of calculation is MODFOLD6, as it integrates different scoring functions and the use of neural networks.

Table 1 Comparison of the input, scoring and output functionalities of the SPSErver and other current servers for the assessment of protein folds

SERVER	REF.	METHOD	INPUT		SCORING		OUTPUT		
			Formats accepted	Accepts multiple structures	Global & Local	Computing time*	Graphical visualiz.	Structure visualiz.	Compare multiple results
SPSErver		Statistical potentials	PDB / CIF	✓	✓	1min	✓	✓	✓
ANOLEA	[14]	Statistical potentials	PDB	✗	✓	<1min	✗	✗	✗
MODFOLD6	[18]	Neural networks	PDB	✓	✓	>1h	✓	✓	✗
ProQ3D	[17]	Neural networks	PDB / CIF	✓	✓	1min	✗	✗	✗
ProSA-web	[6]	Statistical potentials	PDB	✗	✓	<1min	✓	✓	✗
QMEAN	[16]	Statistical potentials and local structural features	PDB / CIF	✓	✓	<1min	✓	✓	✗
Verify 3D	[15]	Local structural features	PDB	✗	✓	<1min	✓	✓	✗
VoroMQA	[8]	Voronoi tessellations	PDB	✓	✓	<1min	✓	✓	✗

*The computing time is approximated using the same structure (1mbn.pdb) and includes the queue waiting time at the moment of submission. We note that the requirement of time may have a strong dependence on the number of users and must not be used to compare the performance

Finally, in terms of intuitive visualization of the results, most web servers offer clear plots for the analysis of local scores. They also provide a tool to visualize the structure, where the residues are colored according to their local score. Still, only the SPSErver provides interactive tools to easily compare the local scores of multiple structures; the local scores can be visualized together in the same plot and smoothed or shifted according to the user's preferences. Additionally, none of the methods reviewed provide tools to score the quality of the interface of PPIs.

Conclusions

The SPSErver facilitates the quality assessment of both protein folds and protein–protein interaction structures in an easy-to-use web server. The quality assessment of the structures is obtained with Split-Statistical Potentials scoring functions that handle several terms associated with the structural local features of the amino acid environments. They are obtained from the analysis of empirical structures: different terms are taken into account such as pairs of interacting residues, solvent accessibility or type of secondary structure. The Split-Statistical Potentials have been tested on the CASP12 dataset and distinguish successfully native structures from wrong decoys. Moreover, the resulting scores are highly correlated with those from reference scoring functions such as PROSA and DOPE. While the other state-of-the-art web servers only show the local scores of the structures in a plot, the SPSErver permits to compare different local score profiles simultaneously. This is done in an interactive plot where the scores can be smoothed or shifted to facilitate the analysis and visualization. Thanks to these analytical tools, we can use the SPSErver to compare the quality of different protein models and protein–protein

interactions, or to understand better the structural effect of a mutation both on the fold and the binding.

Availability and requirements

Project name: SPServer.

Project home page: <https://sbi.upf.edu/spserver>.

Operating system(s): Platform independent.

Programming language: PHP, JavaScript, Python.

Other requirements: Chrome, Safari, Firefox or any other modern browser.

License: Open Source.

Any restrictions to use by non-academics: None.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12859-020-03770-5>.

Additional file 1. Data.

Additional file 2. Figure S1: Identification of steric crashes using GEPOL approach to calculate the surface. The two atoms are represented as light blue and light brown circles. The normal and position vectors are shown both in a case where there is no steric crash (a), and there is a steric crash (b). In the case (a) both vectors form an acute angle (i.e. $< 90^\circ$) while in the case (b) they form an obtuse angle (i.e. $> 90^\circ$), and thus the sign of the two dot products will be negative.

Additional file 3. Figure S2: Residue scores of the native structure of Cysteine synthase A (green), the near-native model (blue) and the wrong model (red). The curves represent the smoothed PAIR scores with a sliding window of value 10.

Additional file 4. Figure S3: Difference between the residue scores of the native structure (reference) and the near-native (blue) and wrong (red) models. The curves represent the smoothed PAIR scores with a sliding window of value 10.

Additional file 5. Figure S4: Local scores map of the interface of the interaction between BAX (Receptor) and BID (Ligand). Large cells are used for local scores (statistic energy) of the wildtype structure and upper (smaller) squares are for the mutant. Energies are shown by colors, from high (red) to low (blue), indicating the range in the label at the bottom. The scores are calculated with the PAIR potential, using a sliding window of 1 to smooth, being the optimal interactions those with most negative energy.

Additional file 6. Figure S5: Mean Pearson correlation values of the comparison between the global scores of the SPServer (ZES3DC and ZPAIR), DOPE and PROSA (Pair Z-score) potentials, and TM, GDT_TS and QCS quality metrics for the structures of CASP12 benchmark. The correlation values are extracted after performing a bootstrapping strategy of 1000 repetitions (described above). The Pearson correlation values of TM score, GDT_TS and QCS are negative because their score is higher when the model is more similar to the native structure (the opposite of the statistical potentials).

Additional file 7. Figure S6: Scatter plots of the global scores of the SPServer potentials ZES3DC (a) and ZPAIR (b) with respect to PROSA (Z-score of Pair potential) for the structures of the CASP12 benchmark.

Additional file 8. Figure S7: Scatter plots of the global scores of the SPServer potentials ZES3DC (a) and ZPAIR (b) with respect to DOPE for the structures of the CASP12 benchmark.

Additional file 9. Figure S8: Scatter plots of the global scores of the SPServer potentials ZES3DC (a) and ZPAIR (b) with respect to GDT_TS for the structures of the CASP12 benchmark.

Additional file 10. Figure S9: Scatter plots of the global scores of the SPServer potentials ZES3DC (a) and ZPAIR (b) with respect to TM score for the structures of the CASP12 benchmark.

Additional file 11. Figure S10: Scatter plots of the global scores of the SPServer potentials ZES3DC (a) and ZPAIR (b) with respect to QCS for the structures of the CASP12 benchmark.

Additional file 12. Figure S11: Histograms showing the residue correlations between the SPServer scoring functions (ES3DC and PAIR) and the PROSA (Pair) and DOPE scoring functions. Each correlation value corresponds to the correlation of all the residue scores of a structure from the CASP12 benchmark.

Additional file 13. Table S1: Global scores of the native structure of Cysteine synthase A and two predicted structural models.

Additional file 14. Table S2: Global scores of the native structure of Cysteine synthase A and two of its models.

Additional file 15. Table S3: Comparison between global and quality metrics for the structures of CASP12 benchmark.

Additional file 16. Table S4: Comparison of local (residue) profiles between SPServer and state-of-art methods DOPE and PROSA for the structures of CASP12 benchmark.

Additional file 17. Table S5: SPServer global Z-scores (ZES3DC, ZPAIR), PROSA (Pair Z-score), DOPE score, GDT_TS, TM score and QCS of the structures of the CASP12 benchmark.

Abbreviations

PPIs: Protein–protein interactions; SPs: Split-statistical potentials; SPServer: Split-statistical potentials server.

Acknowledgments

The authors would like to thank the technical support from GRIB IT team.

Authors' contributions

BO and NFF conceived the project. JAP, MAML and KC designed the web interface. JAP, MAML, JB, OF and BO created the scripts to calculate statistical potentials. GJ created the program to identify steric clashes. AMD wrote the tutorials. JAP, AMD, RM, JB, OF, NFF and BO extensively tested the application. RM tested the scoring functions using the CASP dataset. JAP, AMD, RM and BO designed the case studies. JAP, AMD, RM and BO wrote the manuscript with input from all the authors. All authors read and approved the final manuscript.

Funding

This work was supported by the Spanish Ministry of Science and Innovation (MICINN) [BIO2017-85329-R (co-funded by ERDF,UE)]. BO also acknowledges support from MICINN [ref: MDM-2014-0370]. NFF acknowledges support from [BIO2017-83591-R (co-funded by ERDF, UE)] and [RYC-2015-17519]. We also acknowledge support from the Spanish National Bioinformatics Institute (INB), PRB2-ISCIII and Grants PT13/0001/0023 of the PEI +D+i 2013–2016, funded by ISCIII and co-funded by ERDF of EU. BO and NFF acknowledge Agència de Gestió d'Ajuts Universitaris I de Recerca de la Generalitat de Catalunya, grant SGR17-1020, and the Council for the Catalan Republic contribution to cover the expenses of publication.

Availability of data and materials

The web server can be found at <https://sbi.upf.edu/spserver>. The standalone software can be found at <https://github.com/structuralbioinformatics/SPServer>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Baldo Oliva is member of the Editorial Board of this journal. The rest of authors have no other competing interest.

Author details

¹ Structural Bioinformatics Lab, Department of Experimental and Health Science, Universitat Pompeu Fabra, 08003 Barcelona, Catalonia, Spain. ² Laboratory of Protein Design and Immuno-Engineering, School of Engineering, Ecole Polytechnique Federale de Lausanne, 1015 Lausanne, Vaud, Switzerland. ³ Centre for Molecular Medicine and Therapeutics, Department of Medical Genetics, BC Children's Hospital Research Institute, University of British Columbia, Vancouver, BC V5Z 4H4, Canada. ⁴ Department of Biosciences, U Science Tech, Universitat de Vic-Universitat Central de Catalunya, Vic 08500 Barcelona, Catalonia, Spain. ⁵ Institute of Biological, Environ-Mental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3EB, UK.

Received: 30 April 2020 Accepted: 20 September 2020

Published online: 06 January 2021

References

1. Kryshchuk A, Schwede T, Topf M, Fidelis K, Moutl J. Critical assessment of methods of protein structure prediction (CASP)-Round XIII. *Protein StructFunctBioinform.* 2019;87:1011–20.
2. Won J, Baek M, Monastyrskyy B, Kryshchuk A, Seok C. Assessment of protein model structure accuracy estimation in CASP13: Challenges in the era of deep learning. *Protein StructFunctBioinform.* 2019;87:1351–60. <https://doi.org/10.1002/prot.25804>.
3. Kryshchuk A, Monastyrskyy B, Fidelis K, Schwede T, Tramontano A. Assessment of model accuracy estimations in CASP12. *Protein StructFunctBioinform.* 2018;86:345–60.
4. Cheng J, Choe M, Elofsson A, Han K, Hou J, Maghrabi AHA, et al. Estimation of model accuracy in CASP13. *Protein StructFunctBioinform.* 2019;87:1361–77.
5. Fornes O, Garcia-Garcia J, Bonet J, Oliva B. On the use of knowledge-based potentials for the evaluation of models of protein-protein, protein-DNA, and protein-RNA interactions. In: *Advances in protein chemistry and structural biology.* Elsevier; 2014. p. 77–120.

6. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 2007;35(SUPPL. 2):407–10.
7. Conway P, DiMaio F. Improving hybrid statistical and physical forcefields through local structure enumeration. *Protein Sci.* 2016;25:1525–34.
8. Olechnovič K, Venclovas Č. VoroMQA web server for assessing three-dimensional structures of proteins and protein complexes. *Nucleic Acids Res.* 2019;47:W437–W442442. <https://doi.org/10.1093/nar/gkz367>.
9. Aloy P, Oliva B. Splitting statistical potentials into meaningful scoring functions: testing the prediction of near-native structures from decoy conformations. *BMC StructBiol.* 2009;9:71. <https://doi.org/10.1186/1472-6807-9-71>.
10. Feliu E, Aloy P, Oliva B. On the analysis of protein-protein interactions via knowledge-based potentials for the prediction of protein-protein docking. *Protein Sci.* 2011;20:529–41.
11. Segura J, Marín-López MA, Jones PF, Oliva B, Fernandez-Fuentes N. VORFFIP-driven dock: V-D²OCK, a fast, accurate protein docking strategy. *PLoS ONE.* 2015;10:1–12.
12. Moal IH, Torchala M, Bates PA, Fernández-Recio J. The scoring of poses in protein–protein docking: current capabilities and future directions. *BMC Bioinform.* 2013;14:286. <https://doi.org/10.1186/1471-2105-14-286>.
13. Marín-López MA, Planas-Iglesias J, Aguirre-Plans J, Bonet J, Garcia-Garcia J, Fernandez-Fuentes N, et al. On the mechanisms of protein interactions: predicting their affinity from unbound tertiary structures. *Bioinformatics.* 2017. <https://doi.org/10.1093/bioinformatics/btx616>.
14. Melo F, Devos D, Depiereux E, Feytmans E. ANOLEA: a www server to assess protein structures. *ProclntConfIntell-SystMolBiol.* 1997;5:187–90.
15. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* 1997;277:396–404.
16. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics.* 2011;27:343–50. <https://doi.org/10.1093/bioinformatics/btq662>.
17. Uziela K, Hurtado DM, Shu N, Wallner B, Elofsson A. Pro Q3D: improved model quality assessments using deep learning. *Bioinformatics.* 2017;33:1578–80.
18. Maghrabi AHA, McGuffin LJ. ModFOLD6: an accurate web server for the global and local quality estimation of 3D protein models. *Nucleic Acids Res.* 2017;45:W416–W421421. <https://doi.org/10.1093/nar/gkx332>.
19. Webb B, Sali A. Comparative protein structure modeling using MODELLER. *CurrProtocBioinform.* 2016. <https://doi.org/10.1002/cpps.20>.
20. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015;10:845–58. <https://doi.org/10.1038/nprot.2015.053>.
21. Shi J, Blundell TL, Mizuguchi K. FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J MolBiol.* 2001;310:243–57. <https://doi.org/10.1006/jmbi.2001.4762>.
22. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: Protein structure and function prediction. *Nat Methods.* 2014;12:7–8. <https://doi.org/10.1038/nmeth.3213>.
23. Jones DT, Taylor WR, Thornton JM. A new approach to protein fold recognition. *Nature.* 1992;358:86–9. <https://doi.org/10.1038/358086a0>.
24. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Improved protein structure prediction using potentials from deep learning. *Nature.* 2020;577:706–10. <https://doi.org/10.1038/s41586-019-1923-7>.
25. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, et al. Rosetta3: an object-oriented software suite for the simulation and design of macromolecules. In: Abelson J, et al., editors. *Methods in enzymology*. New York: Academic Press Inc.; 2011. p. 545–574. <https://doi.org/10.1016/B978-0-12-381270-4.00019-6>.
26. Mosca R, Céol A, Aloy P. Interactome3D: adding structural details to protein networks. *Nat Methods.* 2013;10:47–53. <https://doi.org/10.1038/nmeth.2289>.
27. Zhang QC, Petrey D, Garzón JI, Deng L, Honig B. PrePPI: A structure-informed database of protein-protein interactions. *Nucleic Acids Res.* 2013;41:D828–D833833.
28. Meseguer A, Dominguez L, Bota PM, Aguirre-Plans J, Bonet J, Fernandez-Fuentes N, et al. Using collections of structural models to predict changes of binding affinity in protein-protein interactions. *Protein Sci.* 2020. <https://doi.org/10.1002/pro.3930>.
29. Neves MAC, Totrov M, Abagyan R. Docking and scoring with ICM: the benchmarking results and strategies for improvement. *J Comput Aided Mol Des.* 2012;26:675–86. <https://doi.org/10.1007/s10822-012-9547-0>.
30. Jiménez-García B, Pons C, Fernández-Recio J. pyDockWEB: a web server for rigid-body protein–protein docking using electrostatics and desolvation scoring. *Bioinformatics.* 2013;29:1698–9.
31. Gabb HA, Jackson RM, Sternberg MJE. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J MolBiol.* 1997;272:106–20. <https://doi.org/10.1006/jmbi.1997.1203>.
32. Segura J, Marín-López MA, Jones PF, Oliva B, Fernandez-Fuentes N. VORFFIP-driven dock: V-D²OCK, a fast, accurate protein docking strategy. *PLoS ONE.* 2015. <https://doi.org/10.1371/journal.pone.0118107>.
33. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* 2005;33:W363–W367367.
34. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics.* 2014;30:1771–3.
35. Van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastrius PL, Karaca E, et al. The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J MolBiol.* 2016;428:720–5.
36. Krysztafowicz A, Monastyrskyy B, Fidelis K, Schwede T, Tramontano A. Assessment of model accuracy estimations in CASP12. *Proteins StructFunctBioinform.* 2017;86:345–60. <https://doi.org/10.1002/prot.25371>.
37. Eskes R, Desagher S, Antonsson B, Martinou J-C. Bid induces the oligomerization and insertion of bax into the outer mitochondrial membrane. *Mol Cell Biol.* 2000;20:929–35.
38. Meijerink JP, Mensink EJ, Wang K, Sedlak TW, Sløetjes AW, de Witte T, et al. Hematopoietic malignancies demonstrate loss-of-function mutations of BAX. *Blood.* 1998;91:2991–7.

39. López-Blanco JR, Chacón P. KORP: knowledge-based 6D potential for fast protein and loop modeling. *Bioinformatics*. 2019;35:3013–9.
40. Rykunov D, Fiser A. New statistical potential for quality assessment of protein models and a survey of energy functions. *BMC Bioinform*. 2010. <https://doi.org/10.1186/1471-2105-11-128>.
41. Shen M-Y, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci*. 2006;15:2507–24.
42. Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. *Proteins StructFunct Genet*. 2004;57:702–10. <https://doi.org/10.1002/prot.20264>.
43. Zemla A. LGA: a method for finding 3D similarities in protein structures. *Nucleic Acids Res*. 2003;31:3370–4.
44. Cong Q, Kinch LN, Pei J, Shi S, Grishin VN, Li W, et al. An automatic method for CASP9 free modeling structure prediction assessment. *Bioinformatics*. 2011;27:3371–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

