



ELSEVIER

Contents lists available at ScienceDirect

# Biochemistry and Biophysics Reports

journal homepage: [www.elsevier.com/locate/bbrep](http://www.elsevier.com/locate/bbrep)

## Is unphosphorylated Rex, as multifunctional protein of HTLV-1, a fully intrinsically disordered protein? An in silico study



Mitra Kheirabadi <sup>a,\*</sup>, Majid Taghdir <sup>b</sup>

<sup>a</sup> Department of Biology, Faculty of Basic Science, Hakim Sabzevari University, 9617976487 Sabzevar, Iran

<sup>b</sup> Department of Biophysics, Faculty of Biological Sciences, Tarbiat Modares University, Tehran, Iran

### ARTICLE INFO

#### Article history:

Received 9 June 2016

Received in revised form

18 July 2016

Accepted 19 July 2016

Available online 4 August 2016

#### Keywords:

Disordered proteins

HTLV-1

Rex protein

3D model

### ABSTRACT

Intracellular location of a viral unspliced mRNA in host cell is a crucial factor for normal life of the virus. Rex is a nucleo-cytoplasmic shuffling protein of Human T-cell Leukemia Virus-1 (HTLV-1) which has important role in active transport of cargo-containing RNA from nucleus to cytoplasm. Therefore, it plays a crucial role in the disease development by the virus. In spite of its importance, the 3D-structure phosphorylated and unphosphorylated of this protein has not been determined. In this study, first we predicted whether Rex protein is an ordered or disordered protein. In second step protein 3D structure of Rex was obtained. The content of disorder-promoting amino acids, flexibility, hydrophobicity, short linear motifs (SLiMs) and protein binding regions and probability of Rex crystallization were calculated by various In Silico methods. The 3D models of Rex protein were obtained by various In Silico methods, such as homology modeling, threading and ab initio, including; I-TASSER, LOMETS, SPARSKS, ROBBETA and QUARK servers. By comparing and analyzing Qmean, z-scores and energy levels of selected models, the best structures with highest favored region in Ramachandran plot (higher than 90%) was refined with MODREFINER software. In silico analysis of Rex physicochemical properties and also predicted SLiMs and binding regions sites confirms that unphosphorylated Rex protein in HTLV-1 as Rev protein in HIV is wholly disordered protein belongs to the class of intrinsically disordered proteins with extended disorder (native coils, native pre-molten globules).

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

Human T lymphotropic virus type 1 (HTLV-1), a member of Delta-retrovirus family, has single-strand RNA genome and positive polarity [1]. The virus is known as cause of T-cell lymphoma cancer directly and is endemic to various locations around the world, including Africa, South and Central America, the Caribbean region, Asia, and Melanesia [2–4]. HTLV-1 infects human blood cells and causes a variety of diseases in humans, such as adult T-Cell Leukemia (ATL), HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP) and neurological disorders [5–7].

One of the most important conditions for survival of the virus is its ability to move from the nucleus into the cytoplasm, a condition for which REX protein is responsible [8]. REX is a HTLV phosphor-protein, like the REV protein in HIV, and responsible for transporting unspliced mRNA from the nucleus to cytoplasm [9,10]. It has three important functional characteristics in regulating of gene expression of the virus: 1) ability to identify responsive elements in immature RNA before its splicing, 2) ability

to export the processed mRNA from the nucleus through certain pathway, 3) And finally using its regulating domain to monitor its function so that its correct efficient function can be guaranteed [11–18].

So far a large number of research studies have been performed on the functional domain of Rex by using mutation analyses [11–18], which showed the arginine rich area on N-terminal of Rex protein (1–19) that identifies the mRNA virus and attached to it [11–18]. The amino acids 79–99 are responsible for identifying and interacting with human Crm1 protein [11–18].

CRM1 shuttles a very broad range of cargoes including both proteins and RNA between nucleus and cytoplasm using ternary cargo-exportin (CRM1)-Ran GTP complex [19–21]. Moreover, the amino-acids 106–124 and 57–66 take part in the multimerization of Rex on viral mRNA [11–18].

The old idea of the existence of a close relationship between the unique three dimensional structure of a protein and its function has been rejected by the discovery of a new family of proteins in the last decade. According to the old perspectives, a protein needs to adapt to an ordered three dimensional structure for its function. These ordered 3D-structures lead to correct function of proteins. Some

\* Corresponding author.

experimentally evidences in last decade have showed that there are some sections in proteins which play an important role in the protein function but are not compatible with an order-3D-structure. The proteins containing these sections are called intrinsically disordered proteins (IDP) [22–24]. The intrinsically disorder proteins usually serve in the cell signaling functions and also regulating activities. The disorder blocks in IDPs allow the simultaneous interaction with many proteins, and therefore, proteins act as multifunctional proteins in physiological conditions [25,26].

We carried out an in silico analysis of Rex protein in order to determine whether the Rex is an order or disorder protein. There is no experimentally determined 3D structure of Rex Protein available. Therefore, we organized to develop 3D models of Rex Protein by various in silico methods such as homology modeling, threading and ab initio.

### 1.1. Computational methods

Rex Protein sequence with accession No 699504 was obtained from NCBI database at <http://www.ncbi.nlm.nih.gov/protein>.

### 1.2. Disorder prediction

To survey if Rex protein is likely to be an intrinsically disorder protein, MetaDisorder Web service (<http://genesilico.pl/metadisorder/>) and MobiDB web tool (<http://mobidb.bio.unipd.it/>) is used. These web servers combines output of 13 and 10 disorder predictors and generate a consensus prediction, respectively [27–29].

### 1.3. Physicochemical analysis

Rex physicochemical properties were verified using ProtParam (<http://web.expasy.org/protparam/>) [30]. Intrinsically disorder proteins, also have other common characteristics such as high level of flexibility, and frequent hydrophilic amino acids and charged areas; thus these properties were analyzed for Rex protein. Peculiarities of the amino acid composition of Rex protein can be analyzed and visualized by Composition Profiler (<http://www.cprofler.org/>) [31]. To get statistically significant results, a set of sequences of several Rex proteins were selected to analyze (Uniprot ID: P0C205, P0C206 and P0C207). ProtScale (<http://web.expasy.org/protscale/>) as well as Kyte and Doolittle scales were used to determine amino acid hydrophobicity [330]. The average flexibility of Rex was calculated by Kyte and Doolittle scales at (<http://web.expasy.org/protscale/>) and also DynaMine server at (<http://dynamine.ibsquares.be/>) [32]. Charge-hydrophobicity plot and CDF (Cumulative Distribution Function) analysis (both available at <http://www.pondr.com/cgi-bin/PONDR/pondr.cgi>). Use of these tools will allow authors to conclude what type of disorder is present in Rex protein [33–36].

### 1.4. The possibility of Rex crystallization

Either fully or partially disorder proteins have little tendency to crystallization. Therefore, probability of Rex crystallization was checked by using XtalPred server (<http://ffas.burnham.org/XtalPred-cgi/xtal.pl>) [37–40]. “crystallization feasibility” is a single score that illustrate strongly correlation between several protein properties; including protein length, molecular weight, extinction coefficient, instability index, gravy index, and isoelectric point.

### 1.5. SLiMs and protein binding regions in disordered proteins prediction

The content of SLiMs (short linear motifs) in Rex protein predicted by SLiMSearch server at (<http://bioware.ucd.ie/slimsearch2.html>) [41].

ANCHOR algorithm should be utilized to find disorder-based binding sites (<http://anchor.enzim.hu/>) [42,43].

### 1.6. Secondary structure prediction

Secondary structure prediction carried out by PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) [44,45]. PSIPRED is a simple method to predict secondary structure of protein by combining two feed-forward neural networks.

### 1.7. Template selection

PSI-BLAST (Position Specific Iterated- BLAST) against protein data bank (PDB) was performed to identify the most similar structure as template.

### 1.8. 3D Structure prediction

Threading is a fold recognition method to build 3D models for the protein that doesn't have homologs proteins with known structures. 3D models of Rex protein were constructed by threading servers: LOMETS, I-TASSER, and SPARKS. LOMETS at (<http://zhanglab.ccmb.med.umich.edu/LOMETS/>) builds 3D models by collecting high-scoring target-to-template alignments from 9 locally-installed threading programs (FFAS-3D, HHsearch, MUSTER, pGenTHREADER, PPAS, PRC, PROSPECT2, SP3, and SPARKS-X) [46]. I-TASSER server at (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) generates 3D models based on hierarchical method for protein structure and function prediction [47–49]. SPARKSX at (<http://sparks-lab.org/yueyang/server/SPARKS-X/>) constructed 3D models based on application of probabilistic-based matching between predicted primary structural properties of the query and corresponding native properties of templates [50].

Ab initio is a de novo protein structure prediction that builds 3D models of proteins from primary structures in the presence or absence of homologs to query protein. Robetta at ([https://www.google.com/?gws\\_rd=ssl#q=robetta](https://www.google.com/?gws_rd=ssl#q=robetta)) [51] and QUARK at (<http://zhanglab.ccmb.med.umich.edu/QUARK/>) are full-chain protein structure prediction from primary protein structures [52].

### 1.9. 3D structure validation

3D structure validations were done by PROSA, Qmean and Ramachandran web servers. RAMPAGE server at (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) draws a plot to visualize backbone torsional angles  $\psi$  and  $\phi$  of residues in protein structures. It is a universal way to calculate the number of residues in favored, allowed and outlier region in query protein [53]. PROSA web server at (<https://prosa.services.came.sbg.ac.at/prosa.php>) is a frequently employed tools in the validation of protein structures obtained from X-ray analysis, NMR spectroscopy and theoretical calculations [52–55]. Qmean server at (<http://swissmodel.expasy.org/qmean/cgi/index.cgi>) provides access to two scoring functions for the quality estimation of protein structure models which allow to rank a set of models and to identify potentially unreliable region within protein structure [56–58].

### 1.10. 3D structures refinement

Selected models were redefined by Modrefiner server at (<http://zhanglab.ccmb.med.umich.edu/ModRefiner/>). Structural conformational search of proteins is conducted by a combination of physic- and knowledge-based force field [59].

## 2. Results and discussion

Rex is An important post-transcriptional regulatory protein in HTLV-1 life cycle. It has several functional domains to interact with several host cell proteins and virus proteins. Rex is a first mRNA binding protein that selectively binds to unspliced and partially spliced HTLV-1 mRNA, specially binds to the Rex responsive element (RxRE) in the nucleus of host cell to export them into cytoplasm. Viral mRNA codes the structural HTLV-1 proteins. A highly basic N-terminal RNA-binding domain located within amino acids 1–19 is essential for RxRE. This domain also contains a nuclear localization signal (NLS) and binds to p30II, a second HTLV-1 RNA binding protein, to prompt viral replication. The activation domain or NES domain, which encompasses the nuclear export signal (NES), lies between aa 66–118. This domain is critical domain for Rex because Rex interacts with human CRM1 (Exportin-1) through NES. Two multimerization domains, lies between residues 57–66 and 106–124, are located at the –N and –C-terminal ends of NES. The nuclear export signal (NES) domain and multimerization domain are necessary to its nuclear export function. The hCRM1 plays role in multimerization and translocation of Rex. There is a stability domain at the end of C-terminal domain of Rex (lies between residues 170–189). This domain increases the half-life of Rex but it doesn't effect on its function. Monomeric Rex can be exported by hCRM1, but multimerization is necessary for it to interact with unspliced mRNA. Rex is a phosphoprotein which phosphorylation sites are located at Ser170 and 174 in stability domain, Ser97 and Ser106 in NES domain, Thr22, Thr37 and Ser36. Experimental evidences confirmed that phosphorylation of Ser-97 and Thr-174 significantly effect on nuclear export of mRNA HTLV-1 [60].

To continue viral life and protein expression regulation, Rex interacts with several host cellular proteins such as: hCRM1 [61], the heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) [62], the splicing factor SF2 [63], importin  $\beta$  [64], nucleolar protein B-23 [65], Dic2 (Rho-Gap protein) [66], BHLHB2 (a transcription repressor) [67]. In addition, Rex is suspected of interacting with a series of proteins that play crucial roles in mRNA surveillance, nucleocytoplasmic shuttling, tumor growth regulation, and SUMOylation [67].

To investigate the role of the Rex protein in viral life, we analyzed the Rex protein to find out if it is an intrinsically disorder or order protein and also determined its 3D structure using in silico methods.

### 2.1. Disorder structure prediction in Rex

#### 2.1.1. Disorder inducing amino acid structure

Rex Protein amino acid sequence analysis showed that approximately 62.5% of this protein is composed of the so-called disorder inducing amino acids [67–70]. Among them at least 44.4% of primary structure of the Rex protein has strong disorder inducing amino acids such as proline, serine and arginine (Table 1 and Fig. 1). Disorder inducing amino acids such as arginine, glycine, serine, glutamic acid, lysine and proline prevent proteins to fold orderly [67–70]. The high proportion of the disorder inducing amino acids especially proline amino acid (22% prolin in Rex protein) in the sequence of Rex protein suggests that potentially partially or completely disorder area can be found in it. The probability of existing of disorder regions in Rex were predicted using variety servers such as MetaDisorder and also MobiDB (Fig. 1). According to MetaDisorder prediction; the amino acids located in positions 1–32 and 63–189 showed high levels of disorder. In addition, these servers agree in the presence of an order region in 33 to 61 residues. This order region imply to multimerization domains that has important role in interaction with

**Table 1**  
Disorder promoting aminoacids of Rex.

Disorder-promoting amino acids		Order-promoting aminoacids	
		Hydrophobic or bulky	
Polar		Val	2.6%
Ser	15.3%	Phe	2.6%
Pro	21.2%	Trp	1.6%
Glu	2.1%	Tyr	2.6%
Gln	4.8%	Ile	1.6%
Lys	2.1%	Met	1.6%
Arg	7.9%		
Total	61.8%	Total	12.6%

unspliced mRNA. The MobiDB servers predicted that Rex protein is long disorder protein (LD score:74.6% ).

### 2.2. Flexibility

Another important property in disorder proteins is the flexibility. High flexibility levels in secondary and tertiary protein structures lead to fast structural changing, therefore, prevent to form order structures in proteins. And also, these high flexibility levels allow to specifically low affinity interactions [71–73]. Rex Protein flexibility analysis was performed by ExPASy server using an average flexibility scale, DynaMine server (data not be shown) and Composition Profiler server was performed. They showed high flexibility areas that interrupted by short regions with less flexibility levels (Fig. 1). Based on Rex flexibility results, Rex is a protein with high flexibility in most of its regions. Flexibility may affect Rex functions as multifunctional protein which allow the interaction with different cellular targets [71–73].

### 2.3. Hydrophobicity

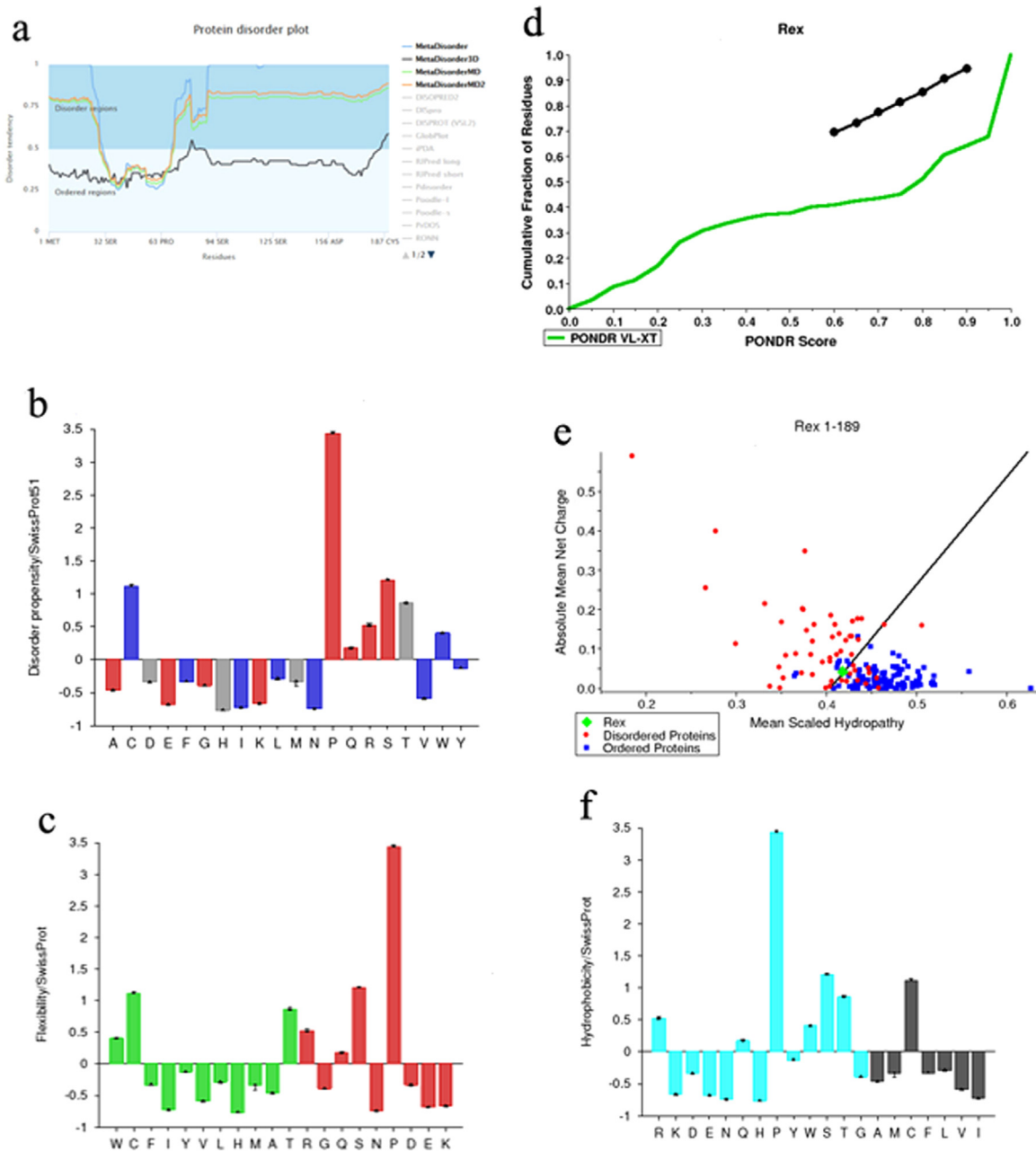
It is clearly obvious that a combination of low hydrophobicity and high protein net charge are two important factors needed to promote disorder regions [74]. Charge-hydropathy plot and CDF analysis, based on Kyte and Doolittle parameter, were revealed that Rex as a wholly disorder protein belongs to the class of IDPs with extended disorder (native coils, native pre-molten globules) (Fig. 1).

### 2.4. Crystallization probability of Rex protein

Crystallization probability was analyzed using XtalPred server. This server calculates probability of crystallization that show the scale from 1 to 5: 1 is for high crystallization probability proteins; while 5 is for very low crystallization probability proteins. A 5 score was obtained for Rex, meaning Rex hardly crystallize.

### 2.5. Short linear motif (SLiMs) and binding regions sites prediction

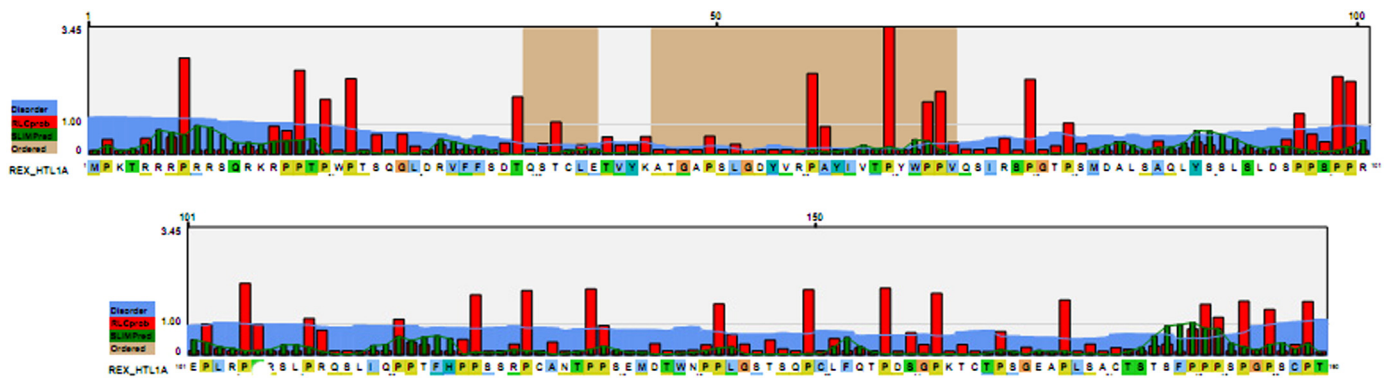
The short linear motifs (SLiMs) or functional microdomains consisting of 3–13 aa play significant role in protein–protein interactions in all aspects of cellular biology. They were first identified in viral life cycle from entry to budding in the host cells [75]. It was discovered later that the viruses actually mimic the functional motifs of cellular proteins to hijack the cellular pathways [75,76]. However, viral mimicry of host SLiMs has not been fully investigated. The SLiMs often take place in intrinsically disordered regions of proteins and such as protein–protein binding (SH3 domain interactions), subcellular targeting (NLS and NES), PTMs (phosphorylation, SUMOylation, and ubiquitination), and cleavage [75].



**Fig. 1.** The primary sequence Rex analysis. (a) Metadisorder server analyzes the probability of disorder structure in primary sequence. The plot shows predicted disorder regions for Rex Protein. These tendency of primary sequence were analyzed using three versions of Metadisorder; blue, green and orange lines. All three, showed similar results that confirmed disorder regions (the values above 0.5) for all amino acids by exception the residues were located at 32–63 positions. (b) Disorder propensity of individual amino acids of Rex protein was calculated using the Composition Profiler server. Disorder-promoting residues (A, R, S, Q, E, G, K, P) are colored red, order-promoting residues (N, C, I, L, F, W, Y, V) are colored blue, and disorder-order neutral residues (D, H, M, T) are colored black. The plot show that Rex protein is enriched by disorder promoting amino acids (P, Q, S and R). (c) The flexibility pattern of individual amino acids of Rex Protein was calculated using the Composition Profiler server. Rigid amino acids place on the left hand side of the plot (green color) and flexible amino acids place on the right hand side of the plot (red color). This plot show that Rex protein is enriched by flexible amino acids (P, S, R and Q). (d) the plot show CDF analysis for Rex protein that protein was calculated using PONDNR server. CDF is used to predict fully disordered or fully ordered proteins. Proteins that have CDF curves above the boundary (boundary in plot is black line with points) were predicted to be ordered, proteins with curves below the boundary are predicted to be disordered, and proteins with curves that crossed the boundary are predicted to be mixture of order and disorder. According to CDF analysis, Rex protein is wholly disorder protein. (e) Charge-hydropathy plot of Rex protein was calculated using PONDNR server. Charge-hydropathy plots compare the absolute, mean net charge and the mean, scaled Kyte-Doolittle hydropathy. The hydropathy measure is scaled between 0 and 1. Disordered proteins are colored red that they separated by linear boundary from ordered proteins are colored blue. According to charge-hydropathy plot, Rex is a disorder protein. (f) Rex hydrophobicity calculated using Composition Profiler server with Kyte and Doolite scale. Hydrophilic amino acids place on the left hand side of the plot (green color) and hydrophobic amino acids place on the right hand side of the plot (black color). This plot show that Rex protein is enriched by hydrophilic amino acids (R, Q, P, W, S and T).

SLiMs prediction for Rex protein was confirmed that there are approximately 12 existing short linear motif. The motifs with high probability place on RNA-binding domain (1–19 aa), NES domain (70–90 aa), the second multimerization domain (106–124 aa) and stability domain (170–189 aa) (Fig. 2). The well-known viral mimicry of HTLV-1 Rex involves NLS and NES in cellular

nucleocytoplasmic shuttling. NES of Rex protein in HTLV acts the same as NES of Rev protein in HIV. NES of HIV as short disorder region interacts weakly to hydrophobic valley on hCRM1 surface. CRM1 exports the variety of proteins and RNAs from nucleus to cytoplasm [77]. The previous experimental results were illustrated that phosphorylation sites on Rex are located at Ser170 and



**Fig. 2.** The SLiMs regions in Rex protein was predicted by using SLiMSearch server. The gray columns mentioned to SLiMs regions. According to prediction, there are 10 SLiMs regions in Rex protein that the motifs with high probability place on RNA-binding domain (1–19 aa), NES domain (70–90 aa), the second multimerization domain (106–124 aa) and stability domain (170–189 aa).

**Table 2**

Models of Rex Protein was calculated by various methods/servers and their evaluation results.

Model/Tool	Method	PROSA z-score	Qmean	Outlier	Allowed region	Favored region
SPARSK-1	Fold recognition	-1.87	0.299	4.3%	8%	87.7%
SPARSK-2	Fold recognition	-1.81	0.195	8.6%	11.2%	80.2%
SPARSK-3	Fold recognition	-0.9	0.225	5.3%	7.5%	87.2%
SPARSK-4	Fold recognition	-0.72	0.176	4.3%	7%	88.8%
SPARSK-5	Fold recognition	-0.41	0.294	5.3%	15%	79.7%
SPARSK-6	Fold recognition	-2.49	0.175	5.3%	5.9%	88.8%
SPARSK-7	Fold recognition	-3.46	0.225	6.4%	9.1%	84.5%
SPARSK-8	Fold recognition	-3.46	0.336	2.7%	11.8%	85.6%
SPARSK-9	Fold recognition	-1.9	0.288	5.3%	15%	79.7%
SPARSK-10	Fold recognition	-1.93	0.277	2.7%	8.6%	88.8%
LEMOT-1	Local meta-threading	-1.82	0.317	6.4%	9.1%	84.5%
LEMOT-2	Local meta-threading	-1.75	0.276	8%	13.9%	78.1%
LEMOT-3	Local meta-threading	-1.59	0.333	6.4%	11.8%	81.8%
LEMOT-4	Local meta-threading	-1.33	0.426	4.3%	7%	88.8%
LEMOT-5	Local meta-threading	-0.59	0.252	5.3%	16%	78.6%
LEMOT-6	<b>Local meta-threading</b>	<b>-1</b>	<b>0.684</b>	<b>0%</b>	<b>2.7%</b>	<b>97.3%</b>
LEMOT-7	<b>Local meta-threading</b>	<b>-0.82</b>	<b>0.685</b>	<b>0.5%</b>	<b>1.6%</b>	<b>97.9%</b>
LEMOT-8	<b>Local meta-threading</b>	<b>0.44</b>	<b>0.586</b>	<b>2.7%</b>	<b>1.1%</b>	<b>96.3%</b>
LEMOT-9	Local meta-threading	0.44	0.120	5.3%	12.8%	81.8%
LEMOT-10	<b>Local meta-threading</b>	<b>-1.6</b>	<b>0.640</b>	<b>0%</b>	<b>2.1%</b>	<b>97.9%</b>
I-TASSER-1	Multiple-threading alignments	-3.29	0.434	9.1%	24.6%	66.3%
I-TASSER-2	Multiple-threading alignments	-2	0.359	24.1%	26.7%	49.2%
I-TASSER-3	Multiple-threading alignments	-0.87	0.304	10.2%	21.4%	68.4%
I-TASSER-4	Multiple-threading alignments	-3.93	0.323	19.3%	34.2%	46.5%
I-TASSER-5	Multiple-threading alignments	-1.47	0.391	6.4%	20.3%	73.3%
QUARK-1	Ab initio	-3.82	0.362	21.9%	13.4%	64.7%
QUARK-2	Ab initio	-1.46	0.240	20.9%	20.3%	58.8%
QUARK-3	Ab initio	-3.19	0.286	22.5%	14.4%	63.1%
QUARK-4	Ab initio	-2.92	0.280	17.1%	13.4%	69.5%
QUARK-5	Ab initio	-3.82	0.341	19.8%	16%	64.2%
QUARK-6	Ab initio	-4.2	0.246	18.2%	18.2%	63.6%
QUARK-7	Ab initio	-3.36	0.254	23%	14.4%	62.6%
QUARK-8	Ab initio	-3.19	0.252	20.9%	17.1%	62%
QUARK-9	Ab initio	-3.11	0.306	19.8%	13.9%	66.3%
QUARK-10	Ab initio	-2.41	0.325	15%	17.1%	67.9%
Robetta-1	Ab initio	-1.8	0.505	3.2%	7.5%	89%
Robetta-2	<b>Ab initio</b>	<b>-1.19</b>	<b>0.500</b>	<b>1.1%</b>	<b>7%</b>	<b>92%</b>
Robetta-3	Ab initio	-2.79	0.455	9.1%	15%	75.9%
Robetta-4	Ab initio	-2.42	0.291	9.6%	12.8%	77.5%
Robetta-5	<b>Ab initio</b>	<b>-2.56</b>	<b>0.523</b>	<b>1.6%</b>	<b>5.3%</b>	<b>93%</b>

174 in stability domain, Ser97 and Ser106 in NES domain, Thr22, Thr37 and Ser36. There is correlation between experimentally detected phosphorylated sites on Rex and predicted short linear motifs. The short linear motifs in intrinsically disorder regions of functional domains of Rex protein are necessary to protein phosphorylation and they can also facilitated protein interactions between Rex and host proteins.

## 2.6. Prediction of protein binding regions in disordered proteins

Many disordered proteins function via binding to a structured partner and undergo a disorder-to-order transition. These binding regions cannot form enough favorable intrachain interactions to fold on their own and are likely to gain stabilizing energy by interacting with a globular protein partner. ANCHOR predicts

**Table 3**

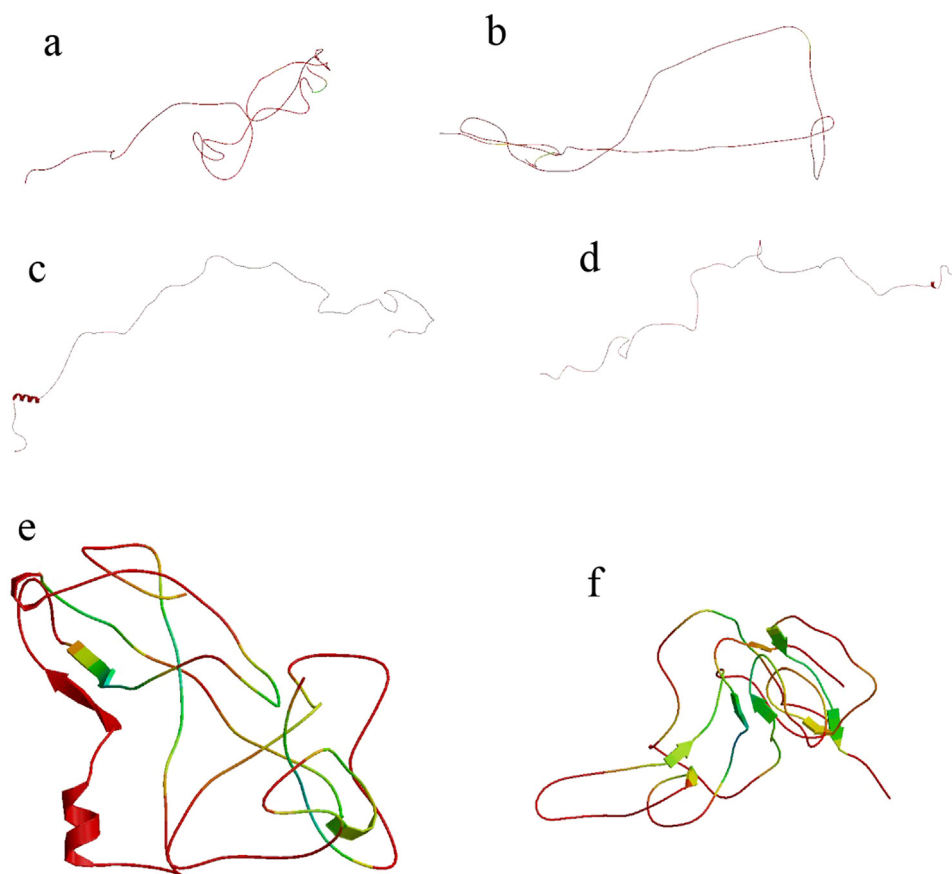
The validated parameters for refined models of Rex Protein.

Model/Tool	Method	PROSA	Qmean	Outlier	Allowed region	Favored Region
LEMOT-6	<b>Local meta-threading</b>	−0.56	<b>0.684</b>	0%	2.7%	<b>97.3%</b>
			0.656	0%	0%	<b>100%</b>
LEMOT-7	<b>Local meta-threading</b>	−0.97	<b>0.685</b>	0.5%	1.6%	<b>97.9%</b>
			0.612	0%	0.5%	<b>99.5%</b>
LEMOT-8	<b>Local meta-threading</b>	−1.59	<b>0.586</b>	2.7%	1.1%	<b>96.3%</b>
			0.572	0.5%	1.1%	<b>98.4%</b>
LEMOT-10	<b>Local meta-threading</b>	−1.29	<b>0.640</b>	0%	2.1%	<b>97.9%</b>
			0.630	0%	0%	<b>100%</b>
Robetta-2	<b>Ab initio</b>	−1.19	<b>0.500</b>	1.1%	7%	<b>92%</b>
			0.492	0.5%	0.5%	<b>98.9%</b>
Robetta-5	<b>Ab initio</b>	−2.56	<b>0.523</b>	1.6%	5.3%	<b>93%</b>
			0.499	0.5%	0.5%	<b>98.9%</b>

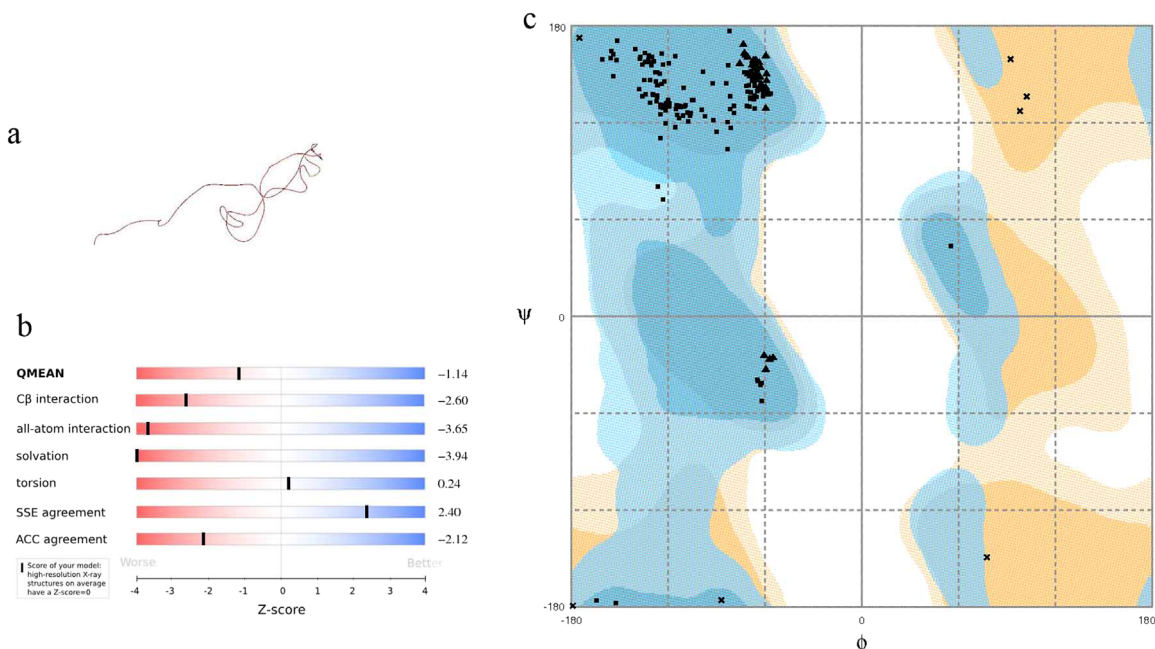
binding regions located in disordered proteins from the amino acid sequence. Based on ANCHOR results, there are five disorder-based binding sites in Rex, ranging in length from 7 to 41 residues. These regions placed on 25–65, 80–92, 108–118, 143–155 and 168–174 amino acids. Three regions of them; NES domain, the second multimerization domain and stability domain is common between SLiMs and binding regions prediction for Rex protein. These results confirmed that three functional domain of Rex have intrinsically disorder protein properties with high probability.

## 2.7. Rex secondary and tertiary structure prediction

Up to now, the 3D-structure of this protein has not been determined, except for a solution NMR structure of the 16-residue fragment of Rex proteins bound to DNA (PDB ID: 1EXY). To get high quality model to Rex, its tertiary structure was modeled using in silico methods including homology modeling, threading and ab initio. Based on PSI-blast results (data not shown), there are no related homologs proteins as a template to construct models by homology modeling. Therefore, threading and ab initio methods were used to construct 3D models for Rex Protein (Table 2). Initially five models built by I-TASSER, ROBETTA and ten models by QUARK, LOMET and SPARSK. All 40 models were used to continue. All models validations were checked by PROSA, Qmean and Ramachandran server (Table 3). First of all, the 6 models with amino acids located in favored region higher than 90% were selected for next step(s). According to Ramachandran plot, Qmean score and z-score, the 4 models constructed by LOMET server (model-6,7,8 and 10) and two models (model-2 and model-5) constructed by Robetta were selected as the best models for Rex Protein (Table 2). The highest percentage for Ramachandran favored region of model-6, model 7, model-8 and model-10 constructed by LOMET server are 97.3, 97.9, 96.3 and 97.9 and model-2 and model-5 are 92 and 93. We selected all of them to refinement (Table 3, Fig. 3). It should be noted that all of the best predicted 3D models of Rex suggest it as a fully disorder protein. Based on the Ramachandran and Qmean and z-score calculated by PROSA server results, model-6 that constructed by LOMET server can be introduced as 3D-structure of Rex protein (Fig. 4).



**Fig. 3.** The best models of Rex protein were predicted with the LOMET and Robetta server. (a) LOMET-6 model, (b) LOMET-7 model, (c) LOMET-8 model, (d) LOMET-10 model, (e) Robetta-2 model and (f) Robetta-5 model. All predicted models by both servers confirmed fully intrinsically disorder for Rex protein.



**Fig. 4.** The 3D model of unphosphorylated Rex protein. (a) the ribbon 3D model of Rex. (b) Qmean parameters for 3D model of Rex protein was calculated using Qmean server. (c) Ramachandran plot of 3D model of Rex to evaluate 3D model. The 100% of its amino acids placed on favored region.

### 3. Conclusion

Disorder proteins are new class proteins that lack order/stable 3D or 2D structures in vivo and in vitro conditions. Partially or fully unfolded proteins play important roles in cell signaling or regulatory phenomena in physiologically condition. They are lack of physiologically stable secondary and tertiary structures which causes multifunctionality. The ability of a protein to fold or not to fold in physiological conditions depends on several factors such as primary structure, a combination of low hydrophobicity and high net charge. These factors are driving forces to protein compaction and also affect the protein crystallization probability. There are many experimental and theoretical evidence that confirmed many viral proteins are either completely disordered or have long functionally disordered regions. These intrinsically disordered regions appear in structural, nonstructural, accessory and regulatory viral proteins [78]. Rev is one of the important regulatory proteins in HIV that its function is the same as Rex protein in HTLV-1. Spectroscopic and hydrodynamic evidence have illustrated that monomeric Rev adopts a molten globule state [79,80]. In silico analysis of Rex physicochemical properties and also predicted SLiMs and binding regions sites and also its 3D-structure model confirms that unphosphorylated Rex protein as Rev protein is wholly disordered protein belongs to the class of intrinsically disordered proteins with extended disorder (native coils, native pre-molten globules).

### Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Appendix A. Transparency document

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.bbrep.2016.07.018>.

### References

- [1] J.H. Richardson, A.J. Edwards, J.K. Cruickshank, P. Rudge, A.G. Dalgleish, In Vivo Cellular Tropism of Human T-Cell Leukemia Virus Type 1, *1*, 64, 11, 1990, 5682–5687.
- [2] G. de Thé, M. Kazanji, An HTLV-I/II vaccine: from animal models to clinical trials? *J. Acquir Immune Defic. Syndr. Hum. Retrovirol.* 13 (1996) 191–198.
- [3] D.U. Gonçalves, F.A. Proietti, J.G.R. Ribas, M.G. Araújo, S.R. Pinheiro, A. C. Guedes, A.B.F. Carneiro-Proietti, Epidemiology, treatment, and prevention of human T-cell leukemia virus type 1-associated diseases, *Clin. Microbiol. Rev.* 23 (3) (2010) 577–589.
- [4] L.J. Roberts, S.E. Huffam, S.F. Walton, B.J. Currie, Crusted scabies: clinical and immunological findings in seventy-eight patients and a review of the literature, *J. Infect.* 50 (2005) 375–381.
- [5] D.U. Gonçalves, F.A. Proietti, E.F. Barbosa-Stancioli, M.L. Martins, J.G. Ribas, O. A. Martins-Filho, A. Teixeira-Carvalho, Peruhype-Magalhães V, A.B. Carneiro-Proietti, HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP) inflammatory network, *Inflamm. Allergy Drug Targets* 7 (2008) 98–107.
- [6] D.U. Gonçalves, F.A. Proietti, E.F. Barbosa-Stancioli, M.L. Martins, J.G. Ribas, O. A. Martins-Filho, A. Teixeira-Carvalho, V. Peruhype-Magalhães, A.B. Carneiro-Proietti, HTLV-1-associated myelopathy/tropical spastic Paraparesis (HAM/TSP) inflammatory network, *Inflamm. Allergy Drug Targets* 7 (2) (2008) 98–107.
- [7] A. Yari, S.A. Rezaee, N. Valizadeh, T. Rajaei, S.M. Jazayeri, M. Soltani, M. Norouzi, Evaluation of HTLV-1 activity in HAM/TSP patients using proviral load and Tax mRNA expression after In Vitro lymphocyte activation, *Iran. J. Basic Med. Sci.* 17 (7) (2014) 531–536.
- [8] I. Younis, P.L. Green, The human T-cell leukemia virus Rex protein, *Front. Biosci.* 10 (2005) 431–445.
- [9] M. Kesic, R. Doueiri, M. Ward, O.J. Semmes, P.L. Green, Phosphorylation regulates human T-cell leukemia virus type 1 Rex function, *Retrovirology* 6 (105) (2009) 1–11.
- [10] S.M. Hanly, L.T. Rimsky, M.H. Malim, J.H. Kim, J. Hauber, M.D. Dodon, S.Y. Le, J. V. Maizel, B.R. Cullen, W.C. Greene, Comparative analysis of the HTLV-I Rex and HIV-1 Rev trans-regulatory proteins and their RNA response elements, *Genes Dev.* 3 (1989) 1534–1544.
- [11] H.P. Bogerd, G.L. Huckaby, Y.F. Ahmed, S.M. Hanly, W.C. Greene, The type 1 human T-cell leukemia virus (HTLV-I) Rex trans-activator binds directly To The HTLV-I Rex and the type 1 human immunodeficiency virus Rev RNA response elements, *Proc. Acad. Sci. USA* 88 (1991) 5704–5708.
- [12] H. Bogerd, W.C. Greene, Dominant negative mutants of human T-cell leukemia virus type 1 Rex and human immunodeficiency virus type 1 Rev fail to multimerize in vivo, *J. Virol.* 67 (1993) 2496–2502.
- [13] T.J. Hope, B.L. Bond, B. McDonald, N.P. Klein, T.G. Parslow, Effector domains of human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type 1 Rex are functionally interchangeable and share an essential peptide motif, *J. Virol.* 65 (1991) 6001–6007.
- [14] H. Siomi, H. Shida, S.H. Nam, T. Nosaka, M. Maki, M. Hatanaka, Sequence requirements for nucleolar localization of human T cell leukemia virus type 1 pX protein, which regulates viral RNA processing, *Cell* 55 (1988) 197–209.

- [15] S.R. Hammes, W.C. Green, Multiple arginine residues within the basic domain of HTLV-1 Rex are required for specific RNA binding and function, *Virology* 193 (1993) 41–49.
- [16] M. Narayan, K. Kusuha, P.L. Green, Phosphorylation of two serine residues regulates human T-cell leukemia virus type 2 Rex function, *J. Virol.* 75 (2001) 8440–8448.
- [17] Weichselbraun I, J. Berger, M. Dobrovnik, et al., Dominant-negative mutants are clustered in adjacent domain of the human T-cell leukemia virus type I Rex protein: implications for transdominance, *J. Virol.* 66 (1992) 4540–4545.
- [18erm] D. Pali, M.H. Malim, The human T-cell leukemia virus type 1 Post-transcriptional trans-activator Rex contains a nuclear export signal, *J. Virol.* 70 (1996) 6442–6445.
- [19] D. Daelemans, S.V. Costes, S. Lockett, G.N. Pavlakis, Kinetic and molecular analysis of nuclear export factor CRM1 association with its cargo in vivo, *Mol. Cell Biol.* 25 (2) (2005) 728–739.
- [20] Y. Yoneda, M. Hieda, E. Nagoshi, Y. Miyamoto, Nucleocytoplasmic Protein transport and recycling of Ran, *Cell Struct. Funct.* 24 (6) (1999) 425–433.
- [21] C. Petosa, G. Schoehn, P. Askjaer, U. Bauer, M. Moulin, U. Steuerwald, M. Soler-López, F. Baudin, I.W. Mattaj, C.W. Müller, Architecture of CRM1/Exportin1 suggests how cooperativity is achieved during formation of a nuclear export complex, *Mol. Cell* 16 (5) (2004) 761–775.
- [22] V.N. Uversky, A.K. Dunker, Understanding protein non-folding, *Biochim. Biophys. Acta* 1804 (2010) 1231–1264.
- [23] H.J. Dyson, P.E. Wright, Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell Biol.* 6 (2005) 197–208.
- [24] P. Tompa, *Structure and Function of Intrinsically Disordered Proteins*. 1st ed., Chapman and Hall/CRC, 2009.
- [25] C.J. Oldfield, A.K. Dunker, Intrinsically disordered proteins and intrinsically disordered protein regions, *Annu. Rev. Biochem.* 83 (2014) 553–584.
- [26] P.E. Wright, H.J. Dyson, Linking folding and binding, *Curr. Opin. Struct. Biol.* 19 (2009) 31–38.
- [27] L.P. Kozłowski, J.M. Bujnicki, MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins, *BMC Bioinform.* 13 (2012) 111.
- [28] E. Potenza, T.D. Domenico, Walsh I, S.C.E. Tosatto, MobiDB 2.0: an improved database of intrinsically disordered and mobile proteins, *Nucleic Acids Res.* (2014), <http://dx.doi.org/10.1093/nar/gku982>.
- [29] T.D. Domenico, I. Walsh, A.J.M. Martin, S.C.E. Tosatto, MobiDB: a comprehensive database of intrinsic protein disorder annotations, *Bioinformatics* 28 (15) (2012) 2080–2081.
- [30] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server; (In), in: John M. Walker (Ed.), *The Proteomics Protocols Handbook*, Humana Press, United States of America, 2005, pp. 571–607.
- [31] V. Vacic, V.N. Uversky, A.K. Dunker, S. Lonardi, Composition Profiler: a tool for discovery and visualization of amino acid composition differences, *BMC Bioinform.* (2007) 008211.
- [32] E. Cilia, R. Pancsa, P. Tompa, T. Lenaerts, W. Vranken, The DynaMine webserver: predicting protein dynamics from sequence, *Nucleic Acid Res.* (2014) 1–7, <http://dx.doi.org/10.1093/nar/gku270>.
- [33] X. Li, P. Romero, M. Rani, A.K. Dunker, Z. Obradovic, Predicting protein disorder for N-, C-, and internal regions, *Genome Inform.* 10 (1999) 30–40.
- [34] P. Romero, Z. Obradovic, X. Li, E. Garner, C. Brown, A.K. Dunker, Sequence complexity of disordered protein, *Proteins: Struct. Funct. Gen* 42 (2001) 38–48.
- [35] P. Romero, Z. Obradovic, A.K. Dunker, Sequence data analysis for long disordered regions prediction in the calcineurin Family, *Genome Inform.* 8 (1997) 110–124.
- [36] P. Romero, Z. Obradovic, C.R. Kissinger, J.E. Villafranca, A.K. Dunker, Identifying Disordered Regions in Proteins from Amino Acid Sequences. Proc. I.E.E.E. International Conference on Neural Networks, 1997, 90–95.
- [37] L. Slabinski, L. Jaroszewski, L. Rychlewski, I.A. Wilson, S.A. Lesley, A. Godzik, XtalPred: a web Server For prediction of protein crystallizability, *Bioinformatics* 23 (24) (2007) 3403–3405.
- [38] L. Slabinski, L. Jaroszewski, A.P.C. Rodrigues, L. Rychlewski, I.A. Wilson, S. A. Lesley, A. Godzik, The challenge of protein structure determination – lessons from structural genomics, *Protein Sci.* 16 (11) (2007) 2472–2482.
- [39] L. Jaroszewski, L. Slabinski, J. Wooley, A.M. Deacon, S.A. Lesley, I.A. Wilson, A. Godzik, Genome pool strategy for structural coverage of protein families, *Structure* 16 (11) (2008) 1659–1667.
- [40] S. Jahandideh, L. Jaroszewski, A. Godzik, Improving the chances of successful protein structure determination with a Random Forest classifier, *Acta Cryst. D Biol. Cryst.* 70 (Pt 3) (2014) 627–6235.
- [41] N.E. Davey, N.J. Haslam, D.C. Shields, R.J. Edwards, SLIMSearch 2.0: biological Context for short linear motifs in proteins, *Nucleic Acids Res.* 39 (Web Server issue) (2011) W56–W60, <http://dx.doi.org/10.1093/nar/gkr402>.
- [42] B. Mészáros, I. Simon, Z. Dosztányi, Prediction of protein binding regions in disordered proteins, *PLoS Comput. Biol.* 5 (5) (2009) e1000376, <http://dx.doi.org/10.1371/journal.pcbi.1000376>.
- [43] Z. Dosztányi, B. Mészáros, I. Simon, ANCHOR: web server for predicting protein binding regions in disordered proteins, *Bioinformatics* 25 (20) (2009) 2745–2746.
- [44] D.W.A. Buchan, F. Minnici, T.C.O. Nugent, K. Bryson, D.T. Jones, Scalable web services for the PSIPRED Protein Analysis Workbench, *Nucleic Acids Res.* 41 (W1) (2013) W340–W348.
- [45] D.T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, *J. Mol. Biol.* 292 (1999) 195–202.
- [46] S. Wu, Y. Zhang, LOMETS: a local meta-threading-server for protein structure prediction, *Nucleic Acids Res.* 35 (2007) 3375–3382.
- [47] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction, *Nat. Methods* 12 (2015) 7–8.
- [48] A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction, *Nat. Protoc.* 5 (2010) 725–738.
- [49] Y. Zhang, I-TASSER server for protein 3D structure prediction, *BMC Bioinform.* 9 (2008) 40.
- [50] Y. Yang, E. Faraggi, H. Zhao, Y. Zhou, Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one dimensional structural properties of the query and corresponding native properties of templates, *Bioinformatics* 27 (2011) 2076–2082.
- [51] D.E. Kim, D. Chivian, D. Baker, Protein structure prediction and analysis using the Robetta server, *Nucleic Acids Res* 32 (Web Server issue) (2004) W526–W531.
- [52] D. Xu, Y. Zhang, Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field, *Proteins* 80 (2012) 1715–1735.
- [53] S.C. Lovell, I.W. Davis, W.B. Arendall III, P.I.W. de Bakker, J.M. Word, M. G. Prisant, J.S. Richardson, D.C. Richardson, Structure validation by Calpha geometry: phi, psi and Cbeta deviation, *Proteins: Struct. Funct. Genet.* 50 (2002) 437–450.
- [54] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res.* 35 (2007) W407–W410.
- [55] M.J. Sippl, Recognition of errors in three-dimensional structures of proteins, *Proteins* 17 (1993) 355–362.
- [56] P. Benkert, M. Künzli, T. Schwede, QMEAN server for protein model quality estimation, *Nucleic Acids Res.* 37 (Web Server issue) (2009) W510–W514.
- [57] P. Benkert, M. Biasini, T. Schwede, Toward the estimation of the absolute quality of individual protein structure models, *Bioinformatics* (2010), <http://dx.doi.org/10.1093/bioinformatics/btq662>.
- [58] P. Benkert, S.C.E. Tosatto, D. Schomburg, QMEAN: a comprehensive scoring function for Model quality assessment, *Protein.: Struct. Funct. Bioinform.* 71 (1) (2008) 261–2277.
- [59] D. Xu, Y. Zhang, Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization, *Biophys. J.* 101 (2011) 2525–22534.
- [60] Y. Adachi, T.D. Copeland, C. Takahashi, T. Nosaka, A. Ahmed, S. Oroszlan, M. Hatanaka, Phosphorylation of the Rex protein of human T-cell leukemia virus type I, *J. Biol. Chem.* 267 (30) (1992) 21977–21981.
- [61] Y. Hakata, M. Yamada, H. Shida, A multifunctional domain in human CRM1 (exportin 1) mediates RanBP3 binding and multimerization of human T-cell leukemia virus type 1 Rex protein, *Mol. Cell Biol.* 23 (23) (2003) 8751–8761.
- [62] E. Kress, H.H. Baydoun, F. Bex, L. Gazzolo, M.D. Dodon, Critical role of hnRNP A1 in HTLV-1 replication in human transformed T lymphocytes, *Retrovirology* 2 (8) (2005) 1–13.
- [63] G.L. Princher, J.G. Julius, S.H. Hughes, D. Derse, Roles of viral and cellular proteins in the expression of alternatively spliced HTLV-1 pX mRNAs, *Virology* 317 (1) (2003) 136–145.
- [64] D. Palmeri, M.H. Malim, Importin beta can mediate the nuclear import of an arginine-rich nuclear localization signal in the absence of importin alpha, *Mol. Cell Biol.* 19 (2) (1999) 1218–1225.
- [65] Y. Adachi, T.D. Copeland, M. Hatanaka, S. Oroszlan, Nucleolar targeting signal of Rex protein of human T-cell leukemia virus type I specifically binds to nucleolar shuttle protein B-23, *J. Biol. Chem.* 268 (19) (1993) 13930–13934.
- [66] N. Simonis, J.F. Rual, I. Lemmens, M. Boxus, T. Hirozane-Kishikawa, J.S. Gatot, A. Dricot, T. Hao, D. Vertommen, S. Legros, S. Daakour, N. Klitgord, M. Martin, J. F. Willaert, F. Dequiedt, V. Navratil, M.E. Cusick, A. Burny, C. Van Lint, D.E. Hill, J. Tavernier, R. Kettmann, M. Vidal, J.C. Twizere, Host-pathogen interactome mapping for HTLV-1 and –2 retroviruses, *Retrovirology* 9 (2012) 26, <http://dx.doi.org/10.1186/1742-4690-9-26>.
- [67] R.M. Williams, Z. Obradovic, Mathura V, W. Braun, E.C. Garner, J. Young, S. Takayama, C.J. Brown, A.K. Dunker, The protein non-folding problem: amino acid determinants of intrinsic order and disorder, *Pac. Symp. Biocomput.* (2001) 89–100.
- [68] P. Romero, Z. Obradovic, X. Li, E.C. Garner, C.J. Brown, A.K. Dunker, Sequence complexity of disordered protein, *Proteins* 42 (2001) 38–48.
- [69] P. Radivojac, L.M. Iakoucheva, C.J. Oldfield, Z. Obradovic, V.N. Uversky, A. K. Dunker, Intrinsic disorder and functional proteomics, *Biophys. J.* 92 (2007) 1439–1456.
- [70] P. Rani, A. Baruah, P. Biswas, Does lack of secondary structure imply intrinsic disorder in proteins? A Sequence, *Anal. Biochim. Biophys. Acta* 1844 (10) (2014) 1827–1834.
- [71] A.K. Dunker, J.D. Lawson, C.J. Brown, R.M. Williams, P. Romero, J.S. Oh, C. J. Oldfield, A.M. Campen, C.M. Ratliff, K.W. Hipps, J. Ausio, M.S. Nissen, R. Reeves, C. Kang, C.R. Kissinger, R.W. Bailey, M.D. Griswold, W. Chiu, E. C. Garner, Z. Obradovic, Intrinsically disordered protein, *J. Mol. Graph Model* 19 (2001) 26–59.
- [72] K. Teillum, J.G. Olsen, B.B. Kragelund, Functional aspects of protein flexibility, *Cell Mol. Life Sci.* 66 (2009) 2231–2247.
- [73] P. Radivojac, Z. Obradovic, D.K. Smith, G. Zhu, S. Vucetic, C.J. Brown, J. D. Lawson, A.K. Dunker, Protein flexibility and intrinsic disorder, *Protein Sci.* 13 (2004) 71–80.



- [74] V.N. Uversky, J.R. Gillespie, A.L. Fink, Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 41 (2000) 415–427.
- [75] N.E. Davey, K. Van Roey, R.J. Weatheritt, G. Toedt, B. Uyar, B. Altenberg, A. Budd, F. Diella, H. Dinkel, T.J. Gibson, Attributes of short linear motifs, *Mol. Biosyst.* 8 (2012) 268–281.
- [76] K. Kadaveru, J. Vyas, M.R. Schiller, Viral infection and human disease – insights from minimotifs, *Front. Biosci.* 13 (2009) 6455–6471.
- [77] T. Güttler, T. Madl, P. Neumann, D. Deichsel, L. Corsini, T. Monecke, R. Ficner, M. Sattler, D. Görlich, NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1, *Nat. Struct. Mol. Biol.* 17 (2010) 1367–1376, doi:10.1038/nsmb.
- [78] B. Xue, D. Blocquel, J. Habchi, A.V. Uversky, L. Kurgan, V.N. Uversky, S. Longhi, Structural Disorder in Viral Proteins, *Chem. Rev.* 114 (13) (2014) 6880–6911, <http://dx.doi.org/10.1021/cr4005692>.
- [79] F. Casu, B.M. Duggan, M. Hennig, The arginine-rich RNA-binding motif of HIV-1 Rev is intrinsically disordered and folds upon RRE binding, *Biophys. J.* 105 (2013) 1004.
- [80] R. Surendran, P. Herman, Z. Cheng, T.J. Daly, J. Ching Lee, HIV Rev self-assembly is linked to a molten-globule to compact structural transition, *Biophys. Chem.* 108 (2004) 101.