

RESEARCH ARTICLE

# The Cold Shock Domain of YB-1 Segregates RNA from DNA by Non-Bonded Interactions

Vladislav Kljashtorny<sup>1,2,3,4\*</sup>, Stanislav Nikonov<sup>1</sup>, Lev Ovchinnikov<sup>1</sup>, Dmitry Lyabin<sup>1</sup>, Nicolas Vodovar<sup>3</sup>, Patrick Curmi<sup>2</sup>, Philippe Manivet<sup>3,4,5\*</sup>

**1** Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290, Russia, **2** Institut National de la Santé et de la Recherche Médicale (INSERM), UMR 829, Laboratoire Structure-Activité des Biomolécules Normales et Pathologiques, Bd François Mitterrand, 91025 Evry Cedex, France, **3** Institut National de la Santé et de la Recherche Médicale (INSERM), UMRS 942, Hôpital Lariboisière, 41 boulevard de la Chapelle, 75475 Paris cedex 10, France, **4** Assistance Publique—Hôpitaux de Paris (APHP), Hôpital Lariboisière, Service de Biochimie et de Biologie Moléculaire, Paris, France, **5** UBCS (Unité de Biologie Clinique Structurale)-Centre de Ressources Biologiques BB-0033-00064, 2 rue Ambroise Paré, 75475 Paris cedex 10, France

\* Current Address: INSERM UMR-S 942 and UBCS-Centre de Ressources Biologiques BB-0033-00064, Hôpital Lariboisière, Paris, France

\* [philippe.manivet@lrp.aphp.fr](mailto:philippe.manivet@lrp.aphp.fr)



CrossMark  
click for updates

## OPEN ACCESS

**Citation:** Kljashtorny V, Nikonov S, Ovchinnikov L, Lyabin D, Vodovar N, Curmi P, et al. (2015) The Cold Shock Domain of YB-1 Segregates RNA from DNA by Non-Bonded Interactions. PLoS ONE 10(7): e0130318. doi:10.1371/journal.pone.0130318

**Editor:** Jean-Luc EPH Darlix, Institut National de la Santé et de la Recherche Médicale, FRANCE

**Received:** December 8, 2014

**Accepted:** May 19, 2015

**Published:** July 6, 2015

**Copyright:** © 2015 Kljashtorny et al. 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 use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** VK receive a PhD grant from the GDR1 "Early Events in Human Diseases (EEHD)" and the Programs on "Molecular and Cellular Biology" and "Basic Sciences to Medicine" from the Presidium of Russian Academy of Sciences. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

The human YB-1 protein plays multiple cellular roles, of which many are dictated by its binding to RNA and DNA through its Cold Shock Domain (CSD). Using molecular dynamics simulation approaches validated by experimental assays, the YB1 CSD was found to interact with nucleic acids in a sequence-dependent manner and with a higher affinity for RNA than DNA. The binding properties of the YB1 CSD were close to those observed for the related bacterial Cold Shock Proteins (CSP), albeit some differences in sequence specificity. The results provide insights in the molecular mechanisms whereby YB-1 interacts with nucleic acids.

## Introduction

The eukaryotic Y-box binding protein 1 (YB-1) is a multifunctional DNA/RNA binding protein that participates in many cellular events through its binding to nucleic acids (NA) and other proteins (reviewed in [1,2,3]).

YB-1 is involved in different aspect of DNA biology. At a transcriptional level, YB-1 binds to the so-called Y-box that is present in the promoter regions of a large number of genes that are involved in cell division, differentiation, apoptosis, and immune and stress responses, and regulates their transcription in a positive or negative fashion (reviewed in [1,3]). YB-1 also displays a high affinity for single-stranded (ss) DNA, DNA damaged regions [4,5], and possesses a RNA and DNA melting and annealing activity [6,7], hence its involvement in DNA repair [8,9].

YB-1 also binds to RNA both in the nucleus and the cytoplasm. YB-1 binds to nascent mRNA [10,11] and is involved in the selection of alternative splice sites [12,13]. In the

cytoplasm, YB-1 is associated with both translated and untranslated mRNAs and interferes with mRNA function and stability in a YB-1/mRNA ratio-dependent manner [14–18]: at low YB-1/mRNA ratio, YB-1 globally activates translation initiation while the saturation of mRNA by YB-1 blocks translation initiation and is accompanied by a complete cessation of protein synthesis [16,17]. An interplay between YB-1 and a series of other protein partners is however necessary to switch mRNA from an untranslated to a translated state. The displacement of YB-1 from mRNA and the stimulation of translation are facilitated by the poly(A) binding protein (PABP) which displays an enhanced affinity for mRNA when interacting with the translation initiation factor eIF4G [19]: the inhibition of translation initiation is caused by the competition between YB-1 and eIF4F for binding to the cap-adjacent region of mRNA [17,18]. Finally, it has been proposed that YB-1 is directly associated with the mRNA cap-structure and that this interaction may be regulated by YB-1 phosphorylation at Ser102 by Akt/PKB kinase [20,21].

YB-1 contains three domains: a N-terminal alanine/proline rich domain (A/P), an intermediate Cold Shock Domain (CSD) which comprises two RNA-binding motifs RNP-1 and RNP-2 that are conserved among eukaryotic YB-1 proteins and prokaryotic CSD proteins (CSP, Fig 1A), and a C-terminal domain which contains alternating positively and negatively charged amino-acid clusters. YB-1 is a rather difficult target for both NMR and X-ray structural methods due to its intrinsic instability that has hampered the solving of its overall structure so far. The only structural information available to date is the NMR structure of the CSD from human YB-1 (named hereafter CSD<sup>YB-1</sup>) [22]. The CSD<sup>YB-1</sup> exhibits a highly conserved beta-barrel structure (Fig 1B), which has been described in many nucleic acids binding proteins [23]. Interestingly, bacterial CSPs consist of a single cold shock domain and are involved to low temperature adaptation [24].

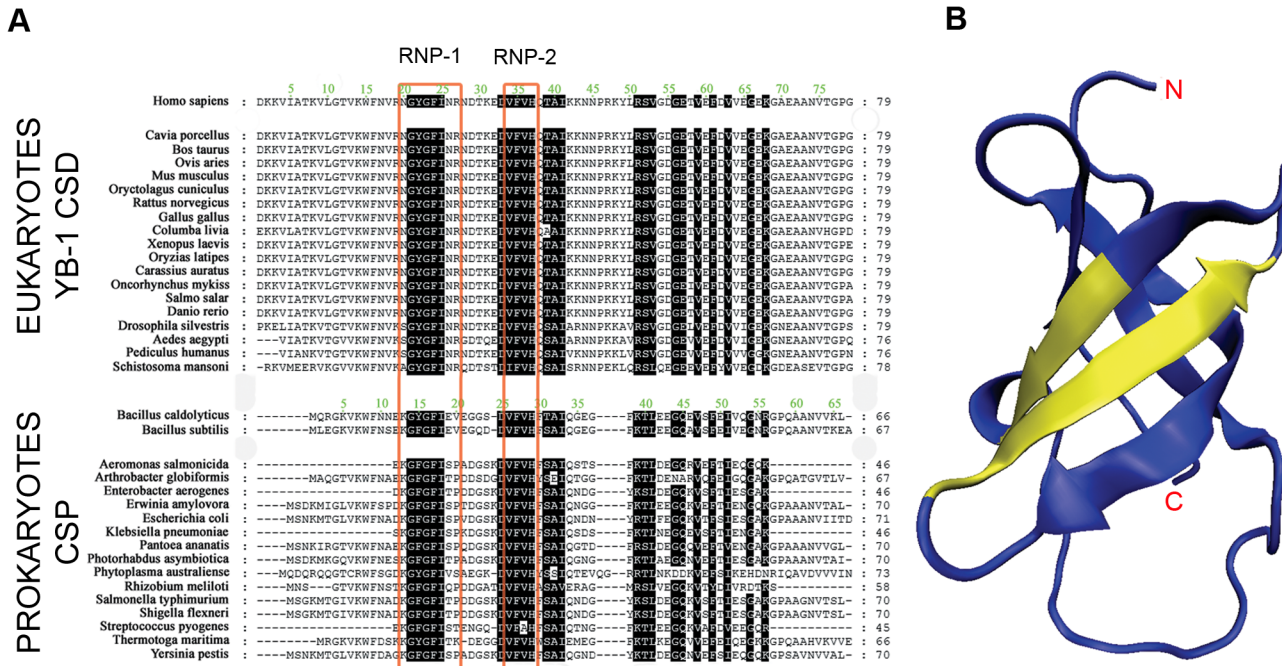
The structural basis for the interaction of YB-1 with nucleic acids remains unclear. A biochemical study showed that the full-length YB-1 exhibit a preferential binding to G-rich DNA [25] and RNA [14] sequences while another supports that, similarly to the bacterial CSPs, the CSD<sup>YB-1</sup> specifically binds to pyrimidine-rich sequences (T or C) [22]. It is worth noting that the affinity of the CSD<sup>YB-1</sup> for oligonucleotides is much lower than that of CSPs [22]. This low affinity hinders biochemical assays and explains the absence of detailed information about the structure of CSD<sup>YB-1</sup>:oligonucleotide complexes. Such structure information is critical to understand the molecular mechanisms of the multiple nucleic acid-dependent functions of YB-1.

We thus propose here, for the first time, to use molecular dynamics simulation (MDS) methods to investigate the interactions of the CSD<sup>YB-1</sup> with short single-stranded desoxyribo- and ribo-oligonucleotides. The results show that the CSD<sup>YB-1</sup> preferentially binds to G-rich sequences as previously stated [14,25], with a higher affinity for RNA than for DNA. Furthermore, the MDS data unraveled the mechanistic basis for the preference of the CSD<sup>YB-1</sup> binding to specific nucleic acid sequences. Finally, the good agreement between the present MDS results and experimental data available to date support MDS as a valuable approach to study the binding properties of nucleic-acid binding proteins.

## Methods

### Generation of the CSD<sup>YB-1</sup>:NA complex.

The NMR structure of the human CSD<sup>YB-1</sup> (PDB-entry 1H95) relaxed after 50 ns of MDS (see below the MDS protocol) was used as a starting model. This structure does not contain any nucleic acid as the interaction between the CSD<sup>YB-1</sup>:NA complex is not stable enough to be solved by NMR [22]. The NMR structure of CSD<sup>YB-1</sup> [22] exhibits a  $\beta$ -barrel fold that is highly similar to the structures of the *Bacillus caldolyticus* Bc-CspB [26] and the *Bacillus Subtilis* Bs-

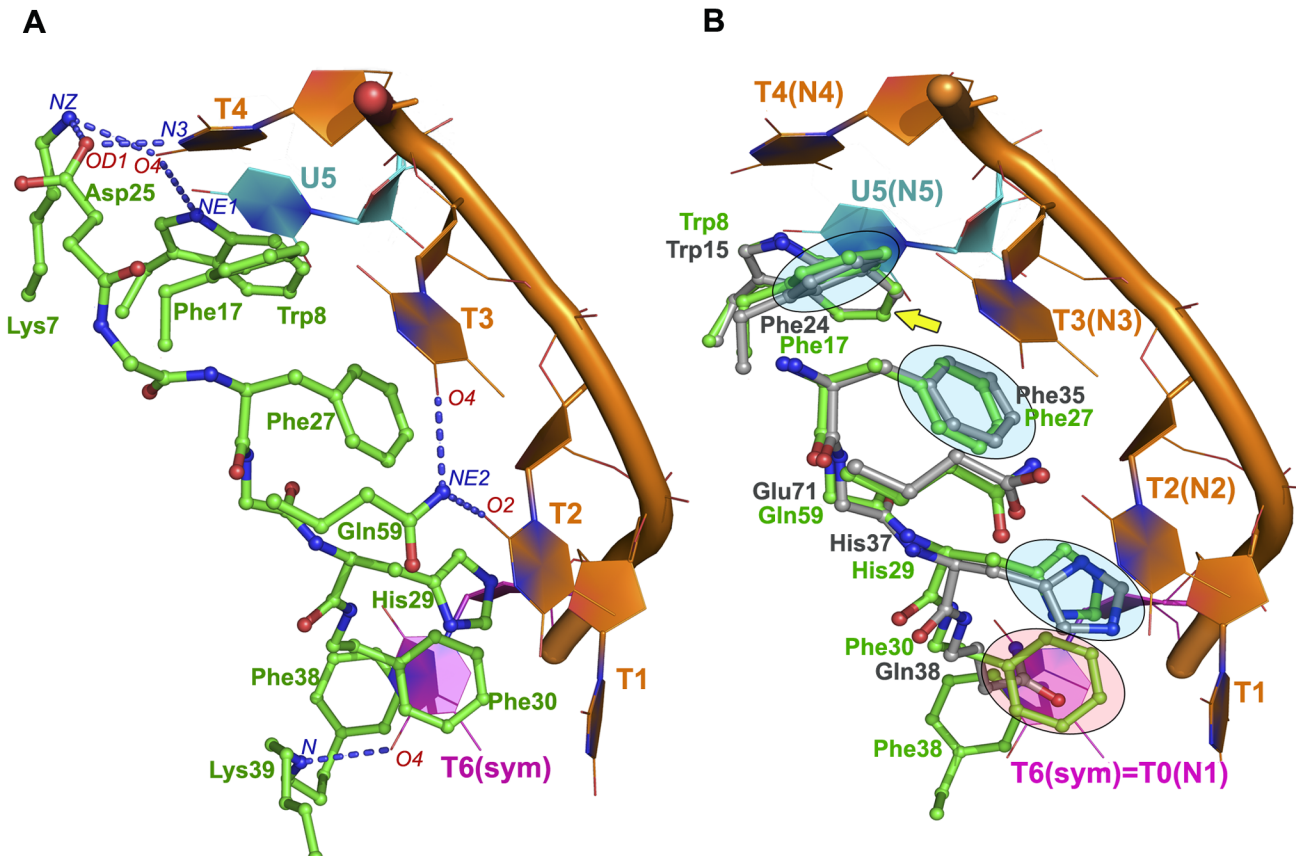


**Fig 1. Sequence alignment of the CSDs from eukaryotic Y-box proteins and prokaryotic CSPs.** **A.** The RNA-binding motifs are boxed in red; highly conserved amino acid residues (identity > 95%) are highlighted in black. **B.** 3D representation of the human CSD<sup>YB-1</sup> structure. The RNP-1 and RNP-2 motifs are shown in yellow; N and C indicate the N- and C-termini of the domain.

doi:10.1371/journal.pone.0130318.g001

CspB [27] CSPs, solved in the presence of single-stranded dT6 oligonucleotides (PDB-entries 2HAX and 2ES2, respectively). Indeed, from a structural standpoint, the nucleic binding sites of the CSD<sup>YB-1</sup> and the aforementioned CSPs are highly conserved as illustrated by an RMSD value of 0.49Å for the 12 C<sub>α</sub>-atoms of the binding site (amino acids 15–18, 26–30, and 59–61 based on Bs-CspB numbering; Fig 1A and 1B). Given the structural similarities observed between the CSD<sup>YB-1</sup> and the bacterial CSP, the coordinates of the oligonucleotide from the Bs-CspB:dT6 were used to build the CSD<sup>YB-1</sup>:DNA complexes used in this study. Of note, the 6<sup>th</sup> residue of the dT6 oligonucleotide from the Bs-CspB:dT6 crystal was bound to symmetrically related Bs-CspB molecule (Fig 2, [26]). Therefore an additional T residue was added at 5'-terminus of the dT6 oligonucleotide to obtain the Bs-CspB:dT7 structure. Using the sugar-phosphate backbone of the oligo(dT) from the Bs-CspB:dT7, each nucleotide was substituted by A, C or G to obtain oligo(dA), oligo(dG) and oligo(dC). One additional nucleotide was added to each extremity, ending up with the CSD<sup>YB-1</sup> bound to nonameric oligonucleotides. The polarity of oligonucleotide obtained in such a way and corresponding to the orientation of dT6 in the bacterial Bs-CspB:dT7 structure is referred as standard or *direct* orientation. In contrast to non-standard or *reversed* orientation where the strand is oriented upside down. To generate such conformation the atoms of the nucleotide bases were aligned to keep the stacking with the amino acid residues of CSD<sup>YB-1</sup> whereas the atoms of sugar and phosphate groups were ignored. Thus, the reversed orientated oligonucleotide preserved the main contacts between nucleotide bases and amino acid residues and only had the flipped polarity of sugar-phosphate backbone.

Similarly, the CSD<sup>YB-1</sup> complexes were generated for ssRNA oligonucleotides. The bacterial numbering scheme of nucleotides bound to the Bs-CspB was used in this work so that the side chains of His29 (His37 in CSD<sup>YB-1</sup>), Phe27 (Phe35 in CSD<sup>YB-1</sup>) and Phe17 (Phe24 in CSD<sup>YB-1</sup>)



**Fig 2. Superposition between the bacterial CSP and the CSD<sup>YB-1</sup> nucleic acid-binding domains.** **A.** Key interactions between the bacterial CSP (green) and an oligo(dT) (orange) obtained from the CSP:oligo(dT) crystal structure. [22] The intermolecular H-bond formed between the protein and the ssDNA are shown as blue dotted lines. **B.** Superposition of the CSD<sup>YB-1</sup> (gray) with the CSP:dT6 complex (green and orange, respectively). The side chains of amino acids involved in stacking with nucleotide bases are highlighted in blue. These three binding sites, that are structurally equivalent between the bacterial CSP and the CSD<sup>YB-1</sup> are labelled N1, N2 and N3. The nucleotide shown in magenta represents the symmetrically related molecule of the complex and shows the additional nucleotide binding site formed by Phe30 and Phe38 (binding site 1, N1), which is only present in the bacterial CSP.

doi:10.1371/journal.pone.0130318.g002

are referred as binding sites 2, 3 and 4, and the corresponding interacting bases of oligonucleotides are referred as N2, N3, N4 (Fig 2B) where N refers to each A, T/U, C and G. When referring to reversed orientation of oligonucleotide we add a little subscript “r” after indication the nucleotide (for instance, dG2r would indicate desoxyriboguanidine bound at the site 2 and having reversed polarity of sugar-phosphate backbone).

More recently, an additional 5<sup>th</sup> binding site has been identified in two crystals of Bs-CspB in complex with short single-stranded RNA molecules (PDB entries 3PF4 and 3PF5 [28]). This 5<sup>th</sup> site that involves the Trp8 residue (Trp15 in CSD<sup>YB-1</sup>) was not observed in the Bs-CspB:DNA complexes, as Trp8 was involved in crystal contacts [26]. The influence of this extra site was evaluated by MD simulations using the coordinate of the polyU oligonucleotide from the Bs-CspB:U6 complex as starting model for MDS. As both starting models gave similar results, only differences between the two setups will be further discussed when applicable.

The comparison of the CSD<sup>YB-1</sup>:dT9 and the Bs-CspB:dT7/U6 structures revealed the absence of some critical amino acid residues in CSD<sup>YB-1</sup> that are involved in the binding of CSPs to nucleic acids. The lack of these amino acids in CSD<sup>YB-1</sup> may contribute to the weak interaction of the CSD<sup>YB-1</sup>:NA complex as compared to that of CSPs:NA and may also explain the difficulty encountered when studying CSD<sup>YB-1</sup>:NA complexes by biochemical and

computational methods such as MDS. MDS experiments consist first in equilibrating the system using position restraints before running simulations under relaxed conditions. In the present study, due to the CSD<sup>YB-1</sup>:NA complexes instability, data acquired starting from the first nanoseconds of the simulation were considered in order to determine the molecular basis for such an instability in CSD<sup>YB-1</sup>:NA complexes.

## Molecular dynamics simulations

MDS were performed with the Gromacs 4.5 software [29] using the Charmm27 force field [30,31]. The starting models of the complexes were placed in an orthogonal water box of 80Å x 60Å x 60Å that contained about 10 000 water molecules of TIP3 type [32]. The water box was generated in the way to keep at least 15Å of bulk water between the surface of the complex and the edge of the water box.

The protonated state for residue was set up as follows: i) the acid functions of acidic residues were deprotonated and carried a -1 charge; ii) the amine functions of basic residues were deprotonated and carried a +1 charge; iii) the phosphate groups of nucleotides were deprotonated and carried a -1 charge. The system was neutralized by the addition of 6 sodium ions to neutralize the -6 charge of the complexes (+3 for the CSD<sup>YB-1</sup> and -9 for an oligonucleotide). Prior to MDS production, the systems were i) minimized to the nearest local minimum using the steepest descent algorithm, ii) equilibrated at a temperature of 310 K for 1 ns with a positional restraints applied to the solute atoms, unless otherwise stated. The systems were then subjected to 50 ns MDS, at constant pressure and temperature (NPT), using a 1 fs integration time-step. The following parameters for NPT MDS were used: i) bond-length were only constrained for water molecules, ii) the long-term non-bonded interactions were evaluated using the PME algorithm, iii) the temperature and pressure were controlled using the velocity rescaling method [33] and the Berendsen algorithm [34], respectively. Coordinates were saved every 0.1 ps. For each the complexes, trajectories were performed 3 times independently.

As mentioned above, prior to simulate the different CSD<sup>YB-1</sup>:NA complexes, the CSD<sup>YB-1</sup> alone was relaxed after 50 ns of MDS by using the same system and MDS protocol as mentioned just here above. Three chlorine ions were used for neutralizing the system.

In order to evaluate the repeatability of our results we performed MDS in triplicate for each CSD<sup>YB-1</sup>:NA complexes by assigning random velocities. No significant differences were obtained for all complexes among the replicates.

## MSD trajectory validation

To validate the trajectories, the NMR data available at BioMagResBank (entry number 4147 [35]) were used. The comparison between experimental and calculated chemical shifts was achieved with Sparta software [36] using C<sub>α</sub>- and C<sub>β</sub>-atoms (S1 Fig). The stereochemical parameters of the models at the different steps of the trajectories were also examined with the Procheck packages [37]. An example of Ramachandran plot is provided in S2 Fig.

## Geometric parameters of the CSD:oligonucleotide interactions

The interactions between the CSD<sup>YB-1</sup> and NA can be described in polar and non-polar terms. The former include H-bonds and electrostatic interactions while the latter involve Van-der-Waals contacts and stacking of aromatic groups. The protein:NA and NA:NA H-bonds as well as the stacking interactions between nucleotides and amino acids were evaluated for each kind of nucleotide at positions 2 to 4 (Fig 2B).

H-bonds were scored at each snapshot of the trajectories when the distance between appropriate H-bond donors and acceptors was less than 3.5Å and the angle acceptor-hydrogen

donor was within 30 degrees. Total number of H-bonds between CSD and each oligonucleotide was counted for each snapshot every 10 ps. For each H-bond, we additionally calculated its occupancy (Occ), *i.e.* the percentage of snapshots along the trajectory where the H-bond was scored. To that end, next expression was applied  $Occ = N/T$ , where N is the number of snapshots where the H-bond was scored, and T is the total number of snapshots.

Another important parameter to take into consideration for H-bonds is their accessibility to the solvent molecules (SA, Å<sup>2</sup>): the less accessible to the solvent an H-bond is, the more it contributes to the stability of the complex [38]. The solvent accessibility for each atom was determined using the *g\_sas* utility from the Gromacs package [29]. Thus, the relative strength of an H-bond depends on both its occupancy and its accessibility to water molecules.

To evaluate stacking interactions, the two following criteria were used: i) the average distance between the geometric centers of the rings of both the base and the aromatic residue during the trajectories, and ii) the angle between the normal of the planes of the rings of the stacked base and the aromatic residue. Well stacked rings show an average distance around 4.2–4.5 Å, and an angle between 0 and 30 degrees throughout the trajectory. These two parameters were also evaluated for each snapshot every 10 ps along the trajectory for each of the complex.

For both H-bond numbers and stacking parameters, a total of 15000 measurements (5000 snapshots from each of 3 trajectories) were collected and pooled for each complex.

## Energetic parameters of the CSD:oligonucleotide interactions

To evaluate binding free energies and to make a ranking for the CSD:NA complexes, the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method [39] was used. According this method the binding free energy ( $\Delta G_{bind}$ ) is calculated as follows:

$$\Delta G_{Bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (1)$$

$$\Delta E_{MM} = \Delta E_{electrostatic} + \Delta E_{vdw} \quad (2)$$

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{SA} \quad (3)$$

where  $\Delta E_{MM}$ ,  $\Delta G_{sol}$  and  $-T\Delta S$  are the changes of the gas phase Molecular Mechanics energy, the solvation free energy, and the conformational entropy upon binding, respectively.  $\Delta E_{MM}$  includes  $\Delta E_{electrostatic}$  (electrostatic), and  $\Delta E_{vdw}$  (van der Waals) energies.  $\Delta G_{sol}$  is the sum of  $\Delta G_{PB}$ , the electrostatic solvation energy (polar contribution), and  $\Delta G_{SA}$ , the non-electrostatic solvation component (non-polar contribution). The polar contribution was calculated here by solving the finite-difference Poisson-Boltzmann equation model using the *PBEQ* module of CHARMM [40]. The grid spacing was set to 0.4 Å, and the longest linear dimension of the grid was extended at least 50% beyond the system; the value of the exterior dielectric constant was set to 80, and the solute dielectric constant was set to each of 1, 2 and 4. The non-polar energy  $\Delta G_{SA}$  was estimated using solvent accessible surface (SAS) as  $0.00542 \times SAS + 0.92 \text{ kcal.mol}^{-1}$  [41]. For the calculation of  $\Delta G_{PB}$  and  $\Delta G_{SA}$ , 50 snapshots (each 1 ns) evenly extracted from the entire MDS trajectories for each complex (0 to 50 ns) were used.  $\Delta E_{MM}$  was calculated based on 5000 snapshots from 0 to 50 ns. The conformational entropy change  $-T\Delta S$  was calculated by quasi-harmonic approach using the *VIBRAN* module of CHARMM [40]. 500000 snapshots extracted from single trajectory were used to evaluate entropic contribution. The calculation for each complex was based on three trajectories and the final value of  $\Delta G_{bind}$  was averaged over 3 measurements.

Calculation methods to measure the binding energy of a system were shown to be dependent on many parameters including the system itself, the simulation protocol, the length of the trajectory, the force-fields, and the solute dielectric constant, and often yield to wrong absolute values [39]. To limit these discrepancies, the relative binding energy of each protein:NA with respect to the most stable of the protein:NA complex (OligoG,  $\Delta\Delta G_{\text{bind}} = \Delta G_{\text{bind}}(\text{oligoN}) - \Delta G_{\text{bind}}(\text{oligoG})$ ) was considered, instead of the absolute binding energy for each protein:NA complex. As for such complex and unstable systems, the sole consideration of relative binding energies may draw biased conclusion, the combination of visual inspection, the averaged parameters of key contacts and MM/PBSA calculations was also used to confidently evaluate the stability of the different protein:NA complexes.

## Biochemical procedures

Recombinant YB-1 and recombinant CSD proteins were purified as described previously [42]. For the electrophoretic mobility shift assay (EMSA), 7 pmol of chemically synthesized RNA (5'-AGGAUGGGUGAGUGAGGUAG-3') or DNA (5'-AGGATGGGTGAGTGAGGTAG-3') oligonucleotides ("DNA-synthesis", Russia) were incubated with increasing amounts of recombinant YB-1 (0.35, 0.7, 1.4, 2.8  $\mu\text{M}$ ) or CSD (3, 6, 12, 24  $\mu\text{M}$ ) in 10  $\mu\text{l}$  reaction buffer (10 mM Hepes-KOH, pH 7.6, 100 mM KCl) at 30°C for 15 minutes. RNA- and DNA-protein complexes were analysed by 8% native polyacrylamide gel electrophoresis (PAGE) in TBE buffer followed by silver staining. Band intensities for free RNA or DNA were determined by gel scanning and quantification using ImageJ software [43].

## Statistical analysis

Data are presented by box plots and include a mean +/- standard deviation or a median and interquartile range, when appropriate. Statistics were performed using the R package (<http://www.r-project.org/>). Analyses of variance (ANOVA) or pairwise t-test were used to compare the data for the different CSD<sup>YB-1</sup>:NA complexes. A two sided type I error of 0.05 was applied with using Bonferroni correction for multiple tests [44] when appropriate.

## Results and Discussion

### The CSD<sup>YB-1</sup> and the bacterial CSPs exhibit different NA binding sites

The CSD<sup>YB-1</sup> and bacterial CSPs belong to the family of the cold shock proteins that exhibit a beta barrel fold that includes two highly conserved RNA-binding motifs, RNP-1 and RNP-2 (Fig 1A), which include most of the residues involved in the interactions with nucleic acids. Both the beta barrel structure and the RNP-1/RNP-2 motifs are highly conserved among CSPs. This structural similarity extends to the CSD domain of YB-1 as illustrated by a 0.49Å RMSD when the structures of the Bc-CspB and the CSD<sup>YB-1</sup> were superposed. Contrary to the bacterial CSPs, the structure of the CSD<sup>YB-1</sup> was not determined in the presence of any nucleic acids. Given the close structural relationship between the CSD<sup>YB-1</sup> and Bs-CspB, their superposition was used to assign the nucleotide binding sites in the CSD<sup>YB-1</sup>. This analysis showed that the nucleotide binding site of the CSD<sup>YB-1</sup> resembles to that of the CSPs, albeit some notable differences: i) the Phe30 residue of the Bs-CspB that is involved in nucleotide stacking is substituted by the non-aromatic Gln38 in CSD<sup>YB-1</sup>, ii) the Gln59 of Bs-CspB that is involved in H-bond with an oligonucleotide base is substituted by Glu71 in CSD<sup>YB-1</sup>; and, iii) most of the bacterial CSPs contain an Phe residue at location 38 which is also stacked to a nucleotide bases, while it is replaced by a Gln in the CSD<sup>YB-1</sup> (Figs 1 and 2; of note, in the crystal structure of the Bs-CspB, this residue interacts with symmetrically related molecule shown in magenta).

Consequently, the bacterial CSPs and the CSD<sup>YB-1</sup> share 3 to 4 structural nucleotide binding sites (Fig 2, N2-N4, N5 was detected only for Bs-CspB:U6 in the crystal structure) while bacterial CSPs exhibit an additional binding site (Fig 2, N1), which is absent in the CSD<sup>YB-1</sup>. Therefore, the nucleotide at position 1 that is located between binding sites N1 and N2 in the bacterial CSPs does not make specific interactions with the binding site N2 of the CSD<sup>YB-1</sup> except by non-specific Van-der-Waals contacts. Altogether, these observations suggest that the differences between the binding sites of bacterial CSPs and the CSD<sup>YB-1</sup> may affect both the binding specificity and the affinity for nucleic acids of CSD<sup>YB-1</sup> as compared to CSPs.

### The CSD<sup>YB-1</sup> and the CSPs exhibit different binding sequence specificity

To evaluate the binding properties of the CSD<sup>YB-1</sup> to nucleic acids, the interaction between the CSD<sup>YB-1</sup> and different homo-oligonucleotides was studied by MDS. Analyses of the MDS trajectories were focused on the three binding sites (N2-N4) that are structurally equivalent in the bacterial CSPs and the CSD<sup>YB-1</sup>. These three sites have been shown to determine the binding specificity of bacterial CSPs to nucleic acids [26,27] and are also expected to govern the binding specificity of CSD<sup>YB-1</sup> to nucleic acids. MDS experiments involving the CSD<sup>YB-1</sup> associated with oligo(dA), oligo(dT), oligo(dC) or oligo(dG) showed that the CSD<sup>YB-1</sup> binds preferentially to the oligo(dG) sequence providing the lowest  $\Delta G$  values and the most stable intermolecular contacts (Table 1). Significant stacking interactions were also observed for oligo(dA) especially in binding sites N2 and N3: the median distance between adenine bases and CSD<sup>YB-1</sup> aromatic residues were 5.1 Å and 4.8 Å, respectively; the median angle between ring planes and CSD<sup>YB-1</sup> aromatic residues were 30° and 23°, respectively (Table 1).

In contrast, the CSD<sup>YB-1</sup>:oligoT and the CSD<sup>YB-1</sup>:oligoC complexes were found significantly less stable with median distances between bases and CSD aromatic residues  $> 7$  Å with only exception of dT at site N2 (distance = 4.8 Å, angle = 26°). In particular, in the CSD<sup>YB-1</sup>:oligo(dC) complex, initial contacts at key positions were lost within the first nanoseconds of the MDS although non-specific interactions prevented the complex to dissociate. All the complexes with oligo(dA), oligo(dT), oligo(dC) had much higher  $\Delta G_{\text{bind}}$  values compared with oligo(dG) which was driven not only by weaker stacking interactions but also by a total number of intermolecular hydrogen bonds. Thus, the average number of H-bonds between CSD and oligo(dG) was 7.1 whereas in the other complexes this number varied from 5.0 to 5.4 (Fig 3). The t-tests for the mean comparisons for the number of intermolecular H-bonds between different complexes revealed the statistically significant difference between all the groups (P-value  $< 0.001$ ). However the difference was only biologically meaningful when compared CSD<sup>YB-1</sup>:oligo(dG) with other complexes. The Cohen's D value [45] for these comparisons ranged between 0.8 and 1 standard deviation with absolute difference of about 2 H-bonds. Such a difference could correspond to a two-folds increase in affinity of CSD<sup>YB-1</sup> for oligo(dG) compared to other oligonucleotides. Among other oligonucleotides oligo(dA) seems to be slightly more preferable ligand for CSD<sup>YB-1</sup> followed by oligo(dT).

Contrary to the CSD<sup>YB-1</sup>, bacterial CSPs exhibit a higher affinity for T-rich sequences [26,27]. The present results show that this specificity is mainly determined by the NH group of Lys39 main chain and the NE2 atom of Gln59 side chain of Bs-CSP, which form strong H-bonds with the O4 and O2 atoms of thymidine at the N1 and N2 sites, respectively (Fig 2A). In the CSD<sup>YB-1</sup>, the N1 site is absent, and Gln59 is substituted by Glu71 at the N2 site. This Gln to Glu substitution at the N2 site completely changes the specificity of the CSD<sup>YB-1</sup> where only a guanine can bind strongly and specifically at this position. This interaction is favored by the formation of strong H-bonds between the side chain of Glu71 and the nitrogen atoms N1 and N2 of dG2 (Fig 4A). These H-bonds had an average occupancy  $> 60\%$ , which was also

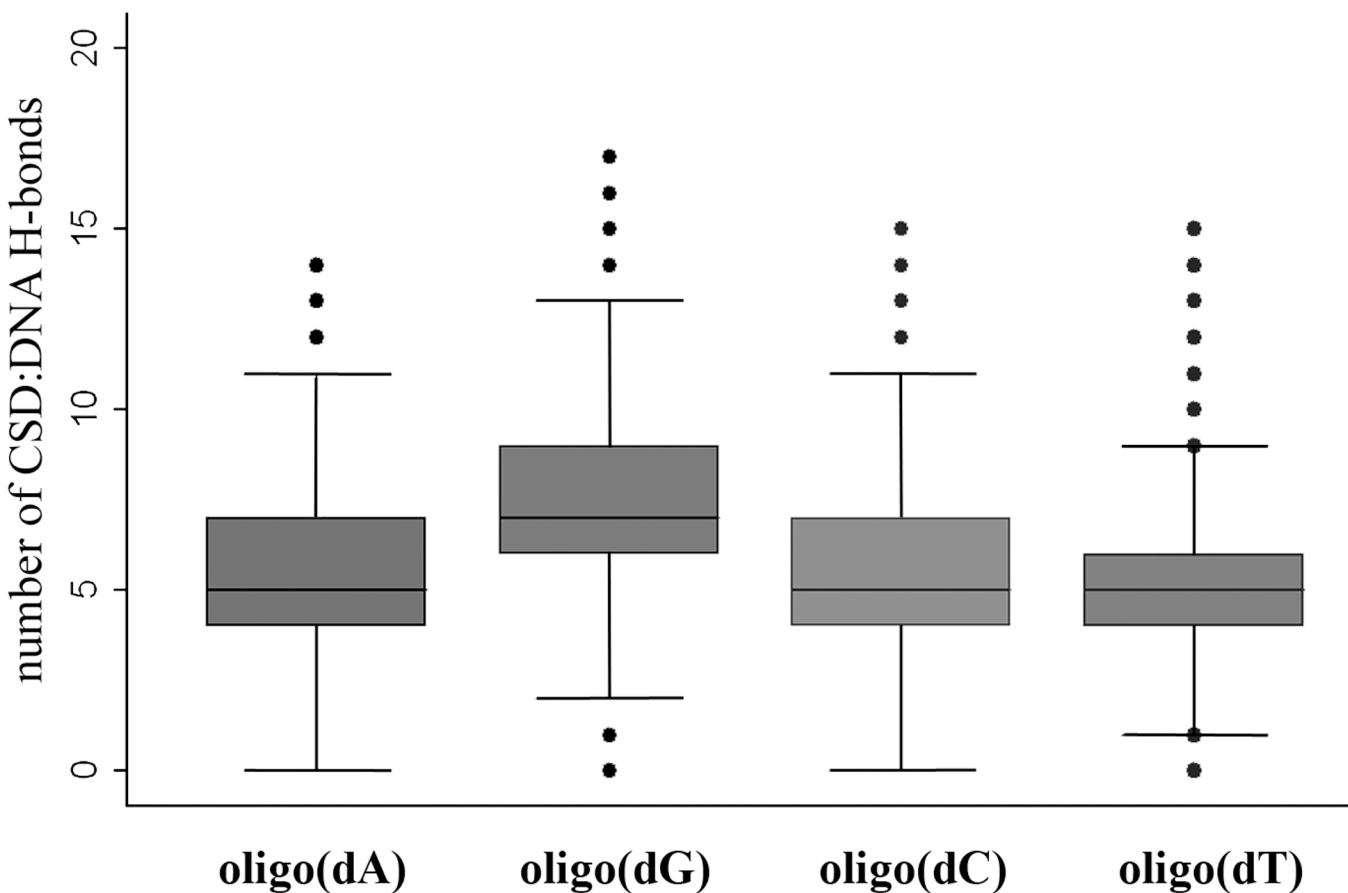


**Table 1. Different types of interactions between CSD<sup>YB-1</sup> and oligonucleotides.** For the parameters of stacking median value during the three trajectories as well as the first and the third quartiles of the data distribution are given. For the H-bond parameters the mean plus or minus a standard deviation are shown.

| Nucleic acid | Nucleotide | Pos2            |            | Pos3            |            | Pos4             |            | Number of NA-CSD H-bonds | Number of NA-NA H-bonds | $\Delta\Delta G_{\text{Bind}}$ kcal/mol |
|--------------|------------|-----------------|------------|-----------------|------------|------------------|------------|--------------------------|-------------------------|---|
|              |            | Distance, Å     | Angle, deg | Distance, Å     | Angle, deg | Distance, Å      | Angle, deg |                          |                         |   |
| DNA          | dG9        | 4.9 (4.6–5.3)   | 27 (18–41) | 4.2 (3.7–7.8)   | 17 (9–56)  | 7.4 (6.9–7.9)    | 79 (70–85) | 7.1 ± 2.1                | 2.0 ± 1.7               | 0                                       |
|              | dA9        | 5.1 (3.9–6.6)   | 30 (15–53) | 4.8 (3.9–7.3)   | 23 (9–66)  | 8.4 (5.5–9.6)    | 71 (53–82) | 5.2 ± 2.2                | 1.6 ± 1.2               | 67                                      |
|              | dT9        | 4.8 (4.0–11.4)  | 26 (15–46) | 6.8 (4.2–7.7)   | 23 (13–37) | 13.1 (6.4–16.5)  | 64 (45–78) | 5.0 ± 2.1                | 1.1 ± 0.8               | 54                                      |
|              | dC9        | 7.6 (3.8–12.8)  | 29 (16–49) | 8.4 (6.8–13.5)  | 54 (30–71) | 12.3 (8.9–16.4)  | 69 (51–81) | 5.4 ± 2.2                | 2.1 ± 1.3               | 23                                      |
|              | dGGTr      | 7.1 (5.3–12.0)  | 48 (33–63) | 5.2 (4.0–7.9)   | 26 (14–50) | 16.0 (14.0–17.6) | 49 (29–69) | 6.4 ± 3.4                | 3.5 ± 2.3               | 35                                      |
| RNA          | G9         | 4.8 (4.5–5.1)   | 20 (14–27) | 4.3 (3.7–7.9)   | 23 (10–66) | 6.6 (5.7–7.7)    | 76 (63–83) | 6.7 ± 2.7                | 7.1 ± 2.2               | 0                                       |
|              | U9         | 8.6 (6.0–12.0)  | 42 (27–61) | 5.4 (3.8–8.0)   | 22 (11–41) | 5.0 (4.0–7.1)    | 40 (21–67) | 3.9 ± 1.7                | 6.4 ± 2.0               | -11                                     |
|              | A9         | 4.3 (4.0–4.7)   | 13 (8–19)  | 6.4 (5.1–8.6)   | 10 (5–38)  | 7.5 (6.9–7.9)    | 80 (71–85) | 4.2 ± 1.9                | 5.1 ± 1.7               | 2                                       |
|              | C9         | 6.5 (3.8–7.3)   | 32 (18–45) | 4.0 (3.8–4.2)   | 13 (8–19)  | 9.2 (6.0–12.0)   | 50 (35–68) | 4.6 ± 1.9                | 7.0 ± 1.9               | 29                                      |
|              | GGUr       | 5.2 (4.7–24.6)  | 44 (34–54) | 4.3 (3.9–20.2)  | 17 (10–38) | 13.5 (10.7–16.2) | 64 (46–77) | 5.9 ± 2.7                | 9.9 ± 2.4               | 8                                       |
|              | U9r        | 12.4 (4.8–22.8) | 50 (26–72) | 18.6 (5.9–22.3) | 52 (32–71) | 19.9 (9.9–23.5)  | 53 (34–72) | 5.6 ± 2.7                | 5.3 ± 1.5               | 30                                      |

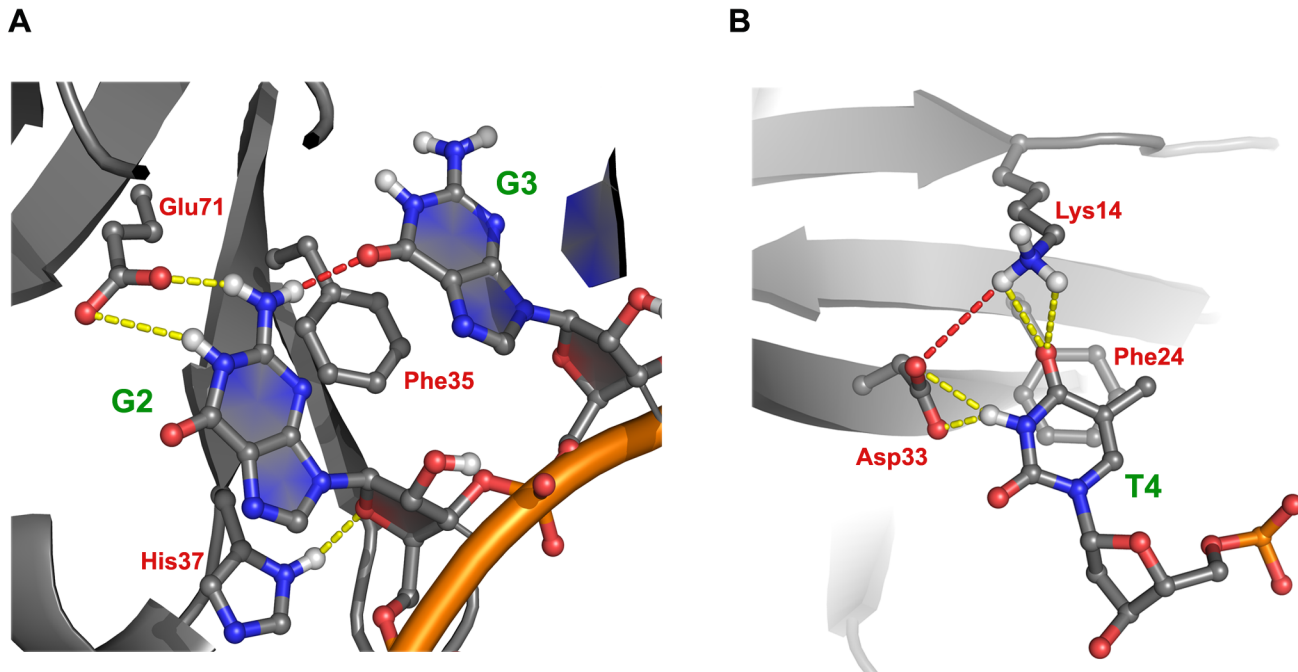
doi:10.1371/journal.pone.0130318.t001

accompanied by a lower solvent accessibility of the respective atoms, strengthening these interactions. In addition, the formation of transient H-bonds between dG2 and dG3 was observed



**Fig 3. Number of intermolecular H-bonds in the complexes of CSD with different desoxyribooligonucleotides.**

doi:10.1371/journal.pone.0130318.g003



**Fig 4. The CSD<sup>YB-1</sup> preferentially binds to GGT sequence.** **A.** Binding of the CSD<sup>YB-1</sup> to oligo(dG) at positions 2 and 3 involves the formation intermolecular H-bonds (yellow dotted lines), including specific H-bonds between G2 and Glu71. In addition, specific intramolecular H-bonds (red dotted lines) between G2 and G3 were also observed. **B.** Binding of the CSD<sup>YB-1</sup> to oligo(dT) at position 4 involves stacking interaction with Phe24 and the formation intermolecular H-bonds (yellow dotted lines) with Asp33 and Lys14. In addition, the interaction of oligo(dT) and the CSD<sup>YB-1</sup> promotes the formation of an intramolecular H-bond between Asp33 and Lys14 (red dotted line).

doi:10.1371/journal.pone.0130318.g004

along the simulation (average occupancy of 25% over the 3 trajectory replicates) which may favor the binding of a guanine at the position N2.

Stable stacking between adenines at positions 2 and 3 and His37 and Phe35 was also detected during the trajectory (Fig 4A). The stacking of dA3 was accompanied with a persistent hydrogen bond between its N6 atom and the main chain oxygen atom of Ala70 (average occupancy of 25% over the 3 trajectory replicates) (S1 Table). However, as mentioned above, the  $\Delta G$  of the CSD<sup>YB-1</sup>:oligo(dA) complex was higher than the  $\Delta G$  of the CSD<sup>YB-1</sup>:oligo(dG) complex, suggesting that the binding of guanines at the N2 and N3 sites is favored over the binding of adenines.

The N4 binding site is almost identical between bacterial CSPs and the CSD<sup>YB-1</sup> (Fig 2B). In agreement with crystallographic data, the MDS results showed some specific H-bonds between the CSP Asp25—Lys7 pair and the polar atoms of the dT4-ring (Fig 2A). In the CSD<sup>YB-1</sup> however, while similar contacts between the Asp33—Lys14 pair and dT4 polar atoms were observed, the interaction was marginally stable (occupancy about 5% over 3 trajectories) compared with the bacterial proteins (Fig 4B).

As in our MDS we observed disruption of most of the complexes, it is logic that CSD<sup>YB-1</sup> involved in the most stable complex should possess lower flexibility compared to less stable ones. To test this hypothesis, we calculated the average atomic fluctuation (RMSF) for the CSD<sup>YB-1</sup> residues and compared them among different complexes (S3 and S4 Figs). The RMSFs for each amino acid residue were calculated based on 50000 snapshots (each 1 ps apart) and averaged over 3 trajectories. Cohen's D effect size for all the comparisons were statistically significant (Table 2). The final rank of the flexibility of CSD<sup>YB-1</sup> amino acid residues can be expressed as follows: Oligo(dG) < Oligo(dA) < Oligo(dT) << Oligo(dC). RMSD calculations

**Table 2. Effect size by Cohen's D for the pairwise comparison of the CSD residues RMSF in the complex with different oligonucleotides.**

| Oligonucleotide | dG9   | dC9    | dA9    | dT9    |
|-----------------|-------|--------|--------|--------|
| dG9             | -     | -1.30* | -0.34* | -0.55* |
| dC9             | 1.30* | -      | 1.00*  | 0.73*  |
| dA9             | 0.34* | -1.00* | -      | -0.24* |
| dT9             | 0.55* | -0.73* | 0.24*  | -      |

\* Indicate statistically significant difference on the significance level of 0.05.

doi:10.1371/journal.pone.0130318.t002

on the same CSD<sup>YB-1</sup>:oligo complexes confirmed the same rank with a net marked higher flexibility for dC oligos (S5 and S6 Figs).

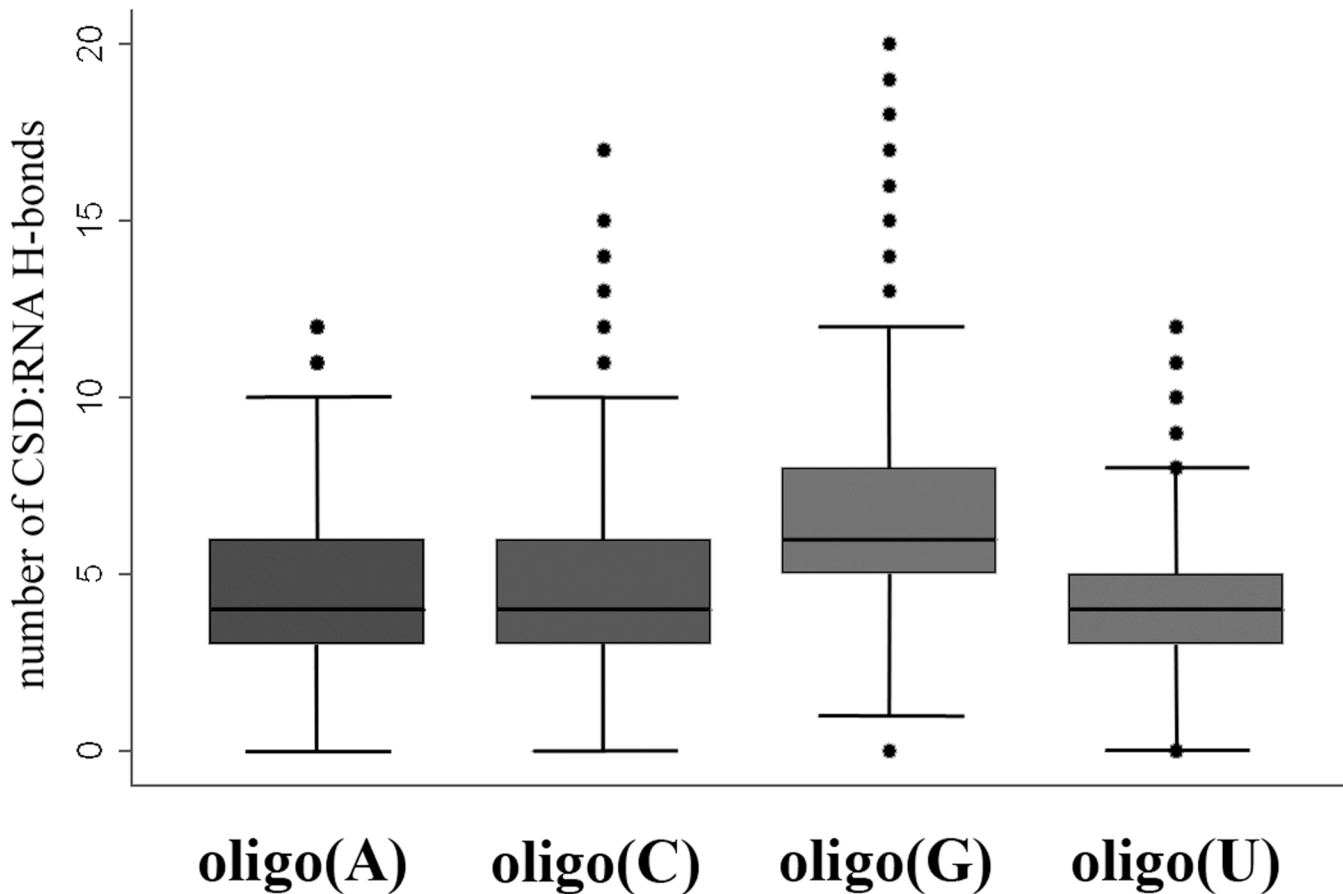
Altogether, these MDS results showed that, in contrast to CSPs that bind preferentially T-rich sequences, the CSD<sup>YB-1</sup> preferentially binds to G-rich sequences. This binding specificity of the CSD<sup>YB-1</sup> is in agreement with the few experimental data available that showed that full-length YB-1 preferentially binds to G-rich sequence [25]. Detailed analysis predicts a GGT triplet to display the highest affinity. These data support a model whereby, the specificity of YB-1 binding to nucleic acid is mainly determined by its CSD domain, while the A/P rich and C-terminal domains may contribute to increase the stability of YB-1:NA complexes [43]. Finally, as previously observed experimentally [22], the present MDS analysis of the CSD<sup>YB-1</sup>:oligo(dG) interaction was found of lower affinity than that measured for the Bs-CSP:oligo(dT) complex. This reduced affinity is due to the loss of the N1 binding site in the CSD<sup>YB-1</sup>, as the N1 site present in the bacterial CSPs promotes strong and specific interactions with nucleic acids.

### The CSD<sup>YB-1</sup> binds RNA with higher affinity than DNA

The relative affinity of the CSD<sup>YB-1</sup> for DNA and RNA was next examined. To this end MD simulations were extended to CSD<sup>YB-1</sup>:oligo(N) complexes. Comparing the RMSD plots for DNA and RNA show clearly that DNA complexes are fluctuating within a larger range of RMSD value pointing a higher flexibility for DNA than RNA (S5 and S6 Figs). As observed for the CSD<sup>YB-1</sup>:DNA interactions, a more stable binding of the CSD<sup>YB-1</sup> was found for the oligo (G) with perfect stacking parameters especially at binding sites 2 and 3 (Table 1) and statistically significant (P-value < 0.001) higher average number of H-bonds between the protein and oligo(G) (Fig 5). Thus, the average number of H-bonds between CSD and oligoG was 6.7 which 0.8 to 1.2 standard deviations higher compared with other CSD<sup>YB-1</sup>:RNA complexes with absolute difference of above 2 H-bonds. Again such a difference could correspond to two folds of higher affinity of CSD for oligoG. The complex CSD<sup>YB-1</sup>:oligoG was used as a reference for  $\Delta\Delta G_{\text{bind}}$  calculation.

Some of the interactions were such unstable that the complex dissociated within the first 5–10 ns of the MDS. Accordingly, we observed increasing fluctuations of the protein residues associated with a decreasing number of CSD:RNA contacts (S3 Fig): the highest fluctuations were observed for the least stable complex CSD<sup>YB-1</sup>:oligoC, although the Cohen's D effect size was moderate (Table 3). Combining the data of different analyses we could rank the oligoribonucleotides by their affinity for CSD<sup>YB-1</sup> as follows: OligoG > OligoA, OligoU > OligoC.

Similarly to what observed with the CSD<sup>YB-1</sup>:oligo(dG) complex, we found specific intramolecular H-bonds between the N2 of G2 and the O6 of G3 (Fig 4A). These H-bonds were found more stable in the case of RNA when compared to DNA in average along the three trajectories. Once again, this specific interaction between guanine at position 2 and 3 suggests a possible cooperative binding of these nucleotides to the N2 and N3 sites.



**Fig 5. Number of intermolecular H-bonds in the complexes of CSD with different ribooligonucleotides.**

doi:10.1371/journal.pone.0130318.g005

Regarding the N4 site, in contrast to DNA, the CSD<sup>YB-1</sup> exhibited more favorable contact with U at position 4 with median distance between U4 and Phe24 of 5.0 Å (Table 1). This stacking is accompanied with the formation of a very strong H-bonds with the Asp33—Lys14 pair (Fig 4B); these H-bonds were significantly less stable when a thymidine (dT4) was stacked at the same position, *i.e.* these two H-bonds were only detected at the beginning of the MDS. Despite this difference at site N4, the binding of G2 and G3 to the amino acids of the N2 and N3 sites did not differ significantly between oligo(dG) and oligo(G).

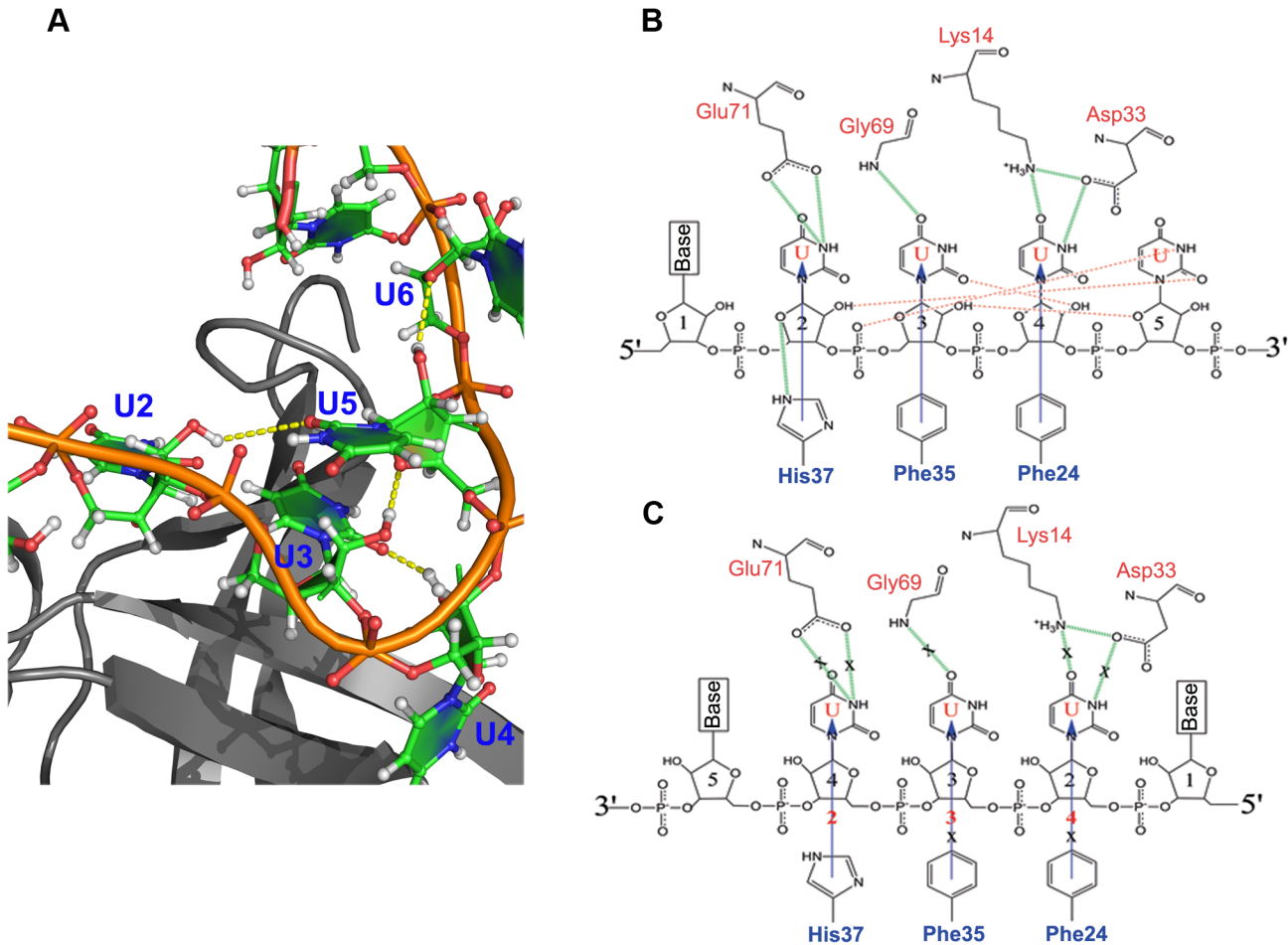
The CSD<sup>YB-1</sup>:oligo(G) interaction was found more stable when compared to the CSD<sup>YB-1</sup>:oligo(dG) interaction (S2 Table). Deoxyribonucleotides differ from ribonucleotides by the lack of an alcohol group at the position 2' of the ribose. We found that this extra-OH groups present

**Table 3. Effect size by Cohen's D for the pairwise comparison of the CSD residues RMSF in the complex with different oligonucleotides.**

| Oligonucleotide | G9    | C9     | A9    | U9     |
|-----------------|-------|--------|-------|--------|
| G9              | -     | -0.23* | 0.03  | -0.27* |
| C9              | 0.23* | -      | 0.27* | -0.04  |
| A9              | -0.03 | -0.27* | -     | -0.30* |
| U9              | 0.27* | 0.04   | 0.30* | -      |

\* Indicate statistically significant difference on the significance level of 0.05.

doi:10.1371/journal.pone.0130318.t003

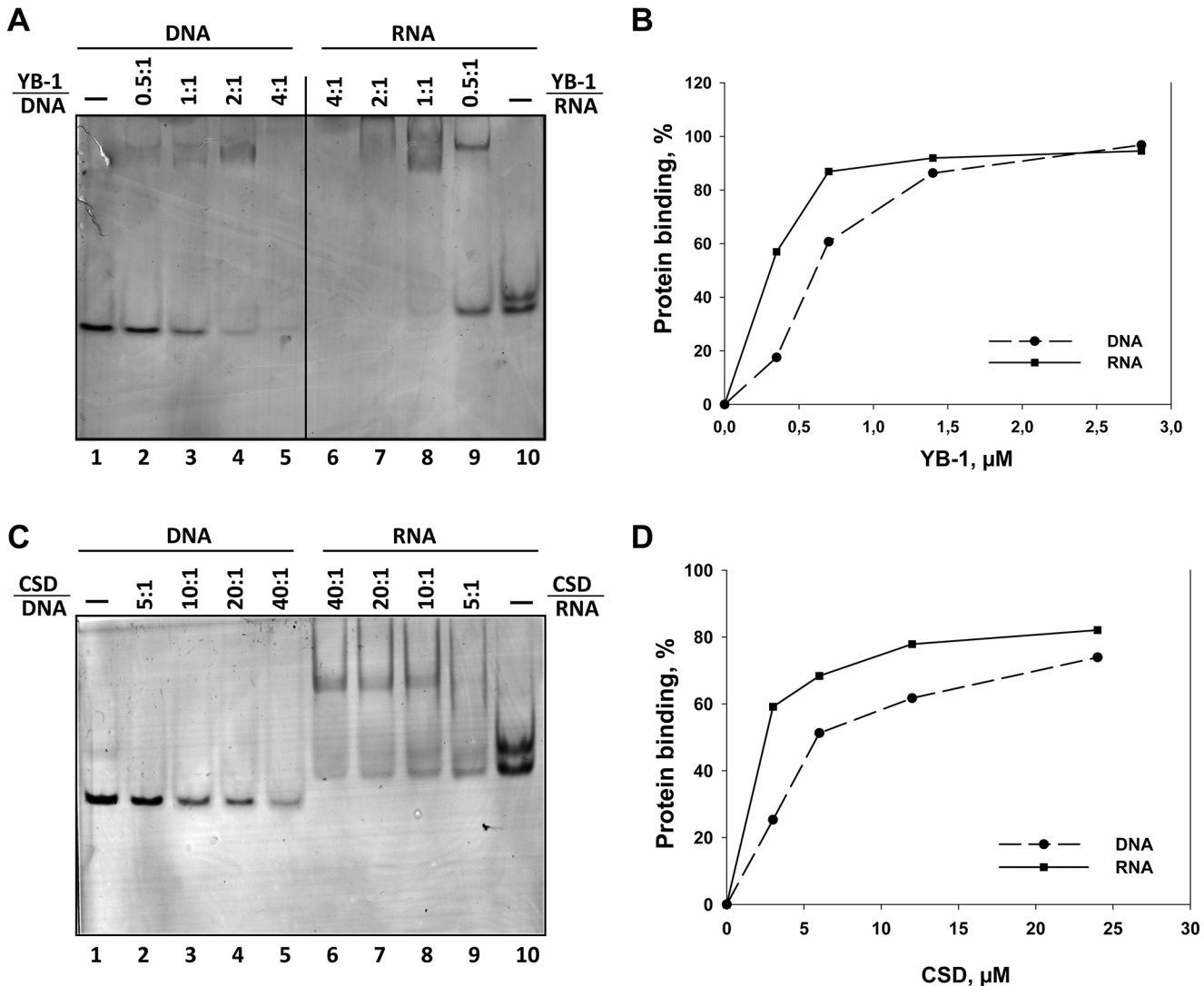


**Fig 6. The CSD<sup>YB-1</sup> binds preferentially to direct-oriented RNA.** **A.** RNA-binding on the surface of CSD (grey) bound to an oligo(U) (green). Intramolecular H-bonds formed between ribonucleotides (yellow dotted lines) stabilise the local conformation of RNA when bound to the CSD. The CSD is shown in grey, the backbone of the oligonucleotide is shown in orange and the bases are shown in green. Nucleotides numbering is indicated. This H-bond network is not formed within DNA molecules. **B** and **C.** Planar representation of the interaction between CSD and direct- (**B**) or reverse-oriented (**C**) oligoribonucleotides. Regardless the orientation of the oligonucleotide, intermolecular H-bonding (green lines) between amino acids and ribonucleotides is identical, with the exception of the H-bond formed between His37 and U2 only observed for direct oligo(U) (**B**). The binding of the CSD<sup>YB-1</sup> to a direct-oriented oligo(U) promotes the formation of intra-oligonucleotide H-bonding (red dashed line) that stabilises the interaction (**B**, H-bonds correspond to **A**). However, this intramolecular network is not formed when CSD binds to the reverse-oriented oligonucleotide (**C**) Consequently, this interaction is less favourable as illustrated by the weakening of the stacking (blue arrow) interaction between amino acids and nucleotides at positions 2 and 3 (marked with X).

doi:10.1371/journal.pone.0130318.g006

in RNA molecules were involved in both CSD<sup>YB-1</sup>:RNA and intra-RNA interactions that were obviously not observed for DNA. In particular, several stable H-bonds were observed between the CSD<sup>YB-1</sup> residues Asn17, Tyr22, Gln38, Glu65, Gly66 and Lys68 and the ribose atoms at position 0, 1, 5, 6 and 8 (**S3 Table**). In addition, specific RNA intra-molecular H-bonds that involved the 2'-OH groups of the ribose of nucleotides 2 to 5 were detected (**Fig 6A**). These RNA-specific bonds should directly contribute to the higher affinity of the CSD<sup>YB-1</sup> to RNA compared to DNA.

To verify this hypothesis, we evaluated the affinity of CSD<sup>YB-1</sup> to G-rich DNA (5'-AGGA TGGGTGAGTGAGGTAG-3') and RNA (5'-AGGAUGGGUGAGUGAGGUAG-3') oligonucleotides *in vitro*. Both oligonucleotides were incubated with increasing amount of recombinant full-length YB-1 or CSD<sup>YB-1</sup>. Quantification of the band intensity of each complex allowed estimating the apparent dissociation constants (**Fig 7** and **S7 Fig**). Both YB-1 (**Fig 7A and 7B**) and



**Fig 7. Determination of the apparent dissociation constants ( $K_a$ ) of the complexes of recombinant YB-1 (A,B) and CSD (C,D) with RNA and DNA.** A and C, Chemically synthesized RNA or DNA fragments of similar sequence and length (7 pmol) were incubated with increasing amounts of recombinant YB-1 (A) or CSD (C) in 10  $\mu\text{l}$  of reaction buffer. The resultant RNA-protein or DNA-protein complexes were analyzed by 8% PAGE followed by silver staining. Protein/NA molar ratio are mentioned at the top of the figures for each experimental condition. B and D, Binding curves were drawn according to quantification of band intensities for free RNA or DNA using ImageJ software. The apparent dissociation constant ( $K_a$ ) was estimated as protein concentration sufficient for 50% saturation. Solid and dotted lines show the RNA-protein and DNA-protein complex, respectively.

doi:10.1371/journal.pone.0130318.g007

CSD<sup>YB-1</sup> (Fig 7C and 7D) bind RNA with more affinity than DNA:  $K_D$  was 0.3  $\mu\text{M}$  and 0.6  $\mu\text{M}$  for the YB-1:RNA and the YB-1:DNA complexes, respectively.  $K_D$  of CSD<sup>YB-1</sup> complexes with RNA and DNA were one order of magnitude lower: 2.5  $\mu\text{M}$  and 6  $\mu\text{M}$ , respectively. These results are in line with our predictions from MDS.

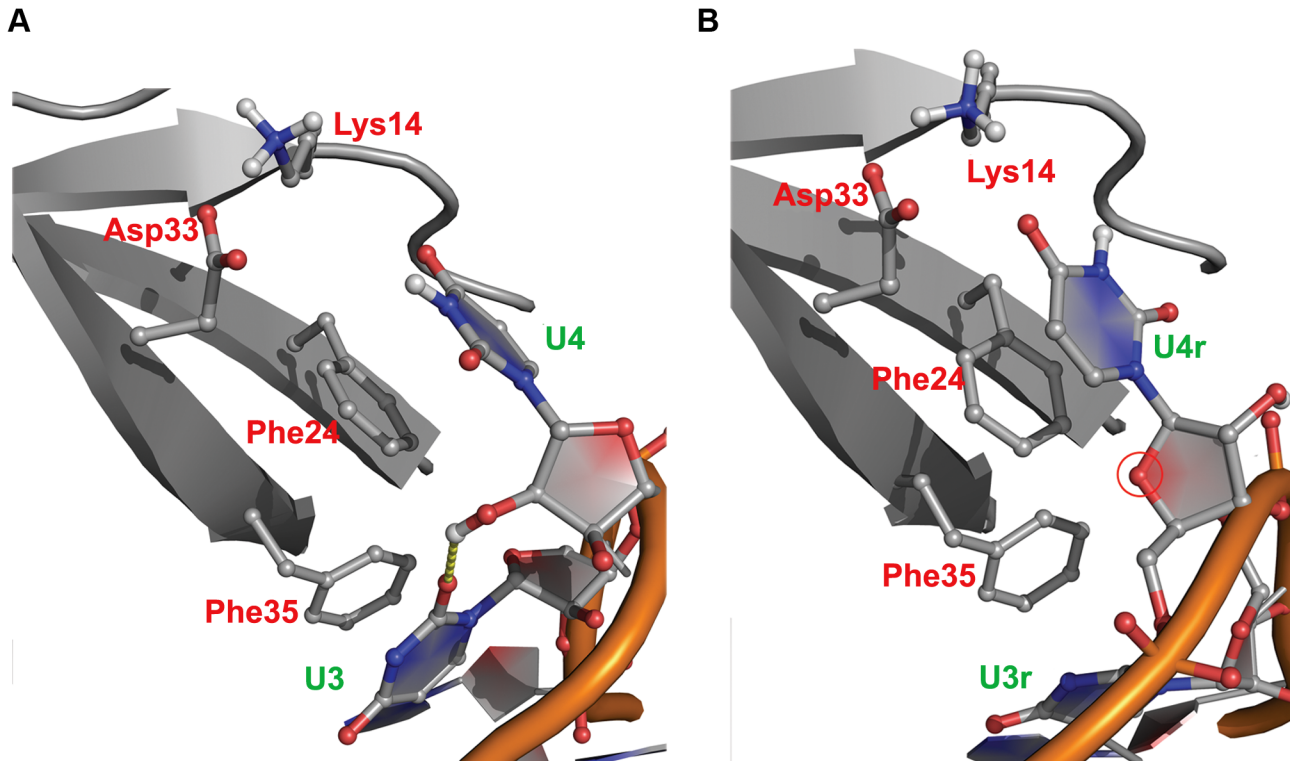
Altogether, these data show that the CSD<sup>YB-1</sup> exhibits a preferential binding to G-rich DNA and RNA sequences. This observation is also in agreement with other experimental data that showed that full-length YB-1 possesses a higher affinity for polyG over polyU and polyA [14]. These data strongly suggest that similarly to what proposed for ssDNA, the CSD<sup>YB-1</sup> is the main contributor that defines the YB-1 binding specificity to RNA. The significant higher affinity of the CSD<sup>YB-1</sup> for RNA can be explained by the presence of numerous H-bonds

between the ribose and the CSD<sup>YB-1</sup>. Finally, the presence of intra-molecular RNA H-bonds may favor the adoption of a local conformation that facilitates either the binding of RNA to the CSD<sup>YB-1</sup> or the stability of the complex. It is worth mentioning that the dramatic change in stability observed between CSD<sup>YB-1</sup>:ssDNA and CSD<sup>YB-1</sup>:ssRNA was not experimentally found in bacterial CSP [28]. Contrary to the CSD<sup>YB-1</sup>:ssRNA complexes where numerous H-bonds within the RNA oligonucleotide stabilized the interaction with the CSD<sup>YB-1</sup>, such intramolecular H-bonds were not observed in the RNA oligonucleotides in complex with bacterial CSP. Therefore, these results re-emphasized a potential role of intramolecular H-bonds in the RNA oligonucleotide as a mechanism that favor the stability of CSD<sup>YB-1</sup>:ssRNA complexes.

### The CSD<sup>YB-1</sup> binds to nucleic acid in a strand orientation-specific fashion

Structural analysis suggests that there is no physical limit for the CSD<sup>YB-1</sup> to bind nucleic acids in either direct or reverse orientations (the “direct” orientation refers here as it appears in the bacterial CSP:oligo(dT) complexes 2ES2 and 2HAX). However it has been shown that all known OB (oligonucleotide/oligosaccharide binding)-fold proteins bind NA in a specific orientation, either direct or reversed [23]. To address the strand orientation specificity of the CSD<sup>YB-1</sup> binding to nucleic acids, the most stable CSD<sup>YB-1</sup>:oligonucleotide complexes in both possible orientations were analyzed by MDS (Fig 6B and 6C). These oligonucleotides contained at positions 2 to 4, a GGT triplet for the DNA oligonucleotides, and a GGU or a UUU triplet for the RNA oligonucleotides. The MDS revealed that most of the inter-molecular contacts (stacking interactions and H-bonds) in the complexes containing an oligonucleotide in reverse orientation were weaker when compared to the same complexes in which the oligonucleotides were in direct orientation (Table 1, S2, S5 and S6 Figs). Accordingly, i) the distance between corresponding nucleotide bases and amino acids residues was significantly higher ( $P < 0.001$ ), ii) the number of intermolecular H-bonds was lower except for CSDYB-1:oligo(Ur), iii) fluctuations of protein residues and oligonucleotide bases were significantly higher ( $P < 0.001$ ), iii) RMSD of protein and nucleotides backbone atoms were higher, when the oligonucleotide was reversed. The increased number of H-bonds for reversed U9 oligonucleotide was due to nonspecific interactions since the initial intermolecular contacts were dissociated very rapidly at the beginning of the trajectories, as illustrated by a distance  $> 12 \text{ \AA}$  median distance between the corresponding residues at the stacking sites over the entire length of the MDS (Table 1).

Analysis of MDS trajectories showed that reversing the oligonucleotide orientation entails unfavorable local interactions between the oligonucleotides tested and the CSD<sup>YB-1</sup>. Although these unfavorable interactions were observed at every position on the oligonucleotide, the most striking effect was observed for the binding site N4. For instance, when the CSD<sup>YB-1</sup> binds to an oligo(U) in direct orientation, the O2' atom of U4 establishes a stable H-bond with the O2 atom of U3 (Fig 8A). In contrast, when the binds an oligo(U) in reverse orientation (rU), the O4' atom of rU4 cannot establish any H-bond neither with atoms from neighboring nucleotides nor with water molecules (Fig 8B). Consequently, the system compensates the loss of H-bonds between the rU4r and the N4 site by interacting with water molecules. Under these conditions, the binding of oligo(U) to the CSD<sup>YB-1</sup> is outcompeted by water molecules, which results in the dissociation of the complex and observed increasing the distance between binding sites and respective nucleotide bases. The strand specificity of the CSD<sup>YB-1</sup>:NA interaction was more prominent when the CSD<sup>YB-1</sup> was bound to reverse oriented G-rich DNA compared with G-rich RNA. In this case, the stacking interactions were weakening more dramatically: the median distance for G-nucleotides at sites 2 was increased from 4.9Å to 7.1Å in the case of DNA and from 4.8Å to 5.2Å in the case of RNA; for the binding site 3 we observed increasing



**Fig 8. Interaction with the CSD<sup>YB-1</sup> and the position 4 of direct- (A) or reverse-oriented (B) RNA oligonucleotides (orange).** A. In the direct orientation, the O2' atom of U4 forms an H-bond with the O2 atom of U3 (yellow dotted line). B. In contrast, the O4' atom of U4r does not form any H-bond, neither with other RNA atom nor with water molecule.

doi:10.1371/journal.pone.0130318.g008

from 4.2Å to 5.2Å for DNA and similar medians of 4.3Å for RNA (Table 1). All increments were statistically significant.

If the estimations of free energies of binding between the different DNA and RNA oligos are in good agreement with experimental data as we were able to demonstrate, difference between direct and reverse orientations of DNA and RNA oligos (especially G-rich ones) are more difficult to interpret due to the limitations of the solvation of Poisson-Boltzmann equations by the MM-PBSA method (S2 Table). This method requires a representative sampling of the conformational space of the solute. In the case of highly flexible molecules like the reverse oriented oligos (see RMSD plots in S4 and S5 Figs), the rapid dissociation from the CSD protein after few picoseconds of MDS, limits the portion of conformational space to sample. In addition, empirical molecular mechanics (MM) force fields like we used in MM-PBSA calculations, cannot describe like quantum chemistry methods, higher order components of non-bonded attractive interactions involved in aromatic molecules stacking. If geometrical parameters of such stacking interactions are well rendered, MM methods used in MM-PBSA calculations penalize the most stable oligos like G-rich ones by underestimating their stability. However, combined with the other geometrical and physical parameters we used in this study, and keeping in mind the above-mentioned limitations, MM-PBSA method remains a good tool for ranking oligo/protein affinities.

Altogether, these data indicate at least one order of magnitude higher affinity of CSD<sup>YB-1</sup> for direct orientated oligonucleotides which is in good agreement with the previously published experimental data for bacterial CSPs [26,27].



## Conclusion

MDS of different complexes between the CSD<sup>YB-1</sup> and different 9-mer oligonucleotides allowed us to document the sequence specificity of the CSD<sup>YB-1</sup> binding to nucleic acids and provided molecular insight into the differential affinity of the CSD<sup>YB-1</sup> for RNA and DNA, in a sequence and strand orientation-specific manner. The differences between the NA binding sites of the CSD<sup>YB-1</sup> and the bacterial CSPs contribute to understand the G-rich binding preference of YB-1 that was previously observed [14,25]. Furthermore, these data support a model whereby the binding specificity of YB-1 can be mainly attributed to its CSD rather than other domains of the protein. Interestingly, the CSD<sup>YB-1</sup> exhibits a significantly lower affinity for nucleic acids than bacterial CSPs. However, contrary to the bacterial proteins that are solely composed of a CSD, YB-1 also contains two additional domains that also contribute to its interaction with nucleic acids [46] and possibly compensate for the relatively low affinity of the CSD for nucleic acids. Unfortunately, the absence of available structure for the other domains of YB-1 prevents the full-length protein to be studied by MDS. MDS also unraveled the role of the 2'-OH of the nucleic acid sugar moiety in stabilizing the interaction between the CSD<sup>YB-1</sup> and RNA molecules. The presence of the 2'-OH in ribose contributed in the establishment of CSD<sup>YB-1</sup>:RNA H-bonds along with the formation of intra-molecular H-bonds that could maintain RNA molecules in a more favorable conformation than DNA upon binding to the CSD<sup>YB-1</sup>. In conclusion, MDS of the CSD<sup>YB-1</sup> in interaction with nucleic acids appears as a valuable model to predict the binding properties of YB-1 for nucleic acids and understand its complex involvement in transcription and translation.

## Supporting Information

**S1 Fig. Correlation plots between experimental and calculated NMR chemical shifts obtained from different structures for <sup>13</sup>C<sub>α</sub> (left) and <sup>13</sup>C<sub>β</sub> (right) atoms of the human CSD<sup>YB-1</sup>: initial structure available in PDB (top), the last point structure after simulation (bottom).**

(TIF)

**S2 Fig. Ramachandran plots of human the CSD<sup>YB-1</sup> allow to check the stereochemical quality of the models at different stages along MD trajectories: initial structure (left) and the structure after 50 ns simulation (right).** The percentage of the amino-acid residues localized in the most preferable regions of the diagram is shown for each model.

(TIF)

**S3 Fig. Atomic fluctuations (RMSF) of the protein C<sub>α</sub>-atoms along the MD trajectories starting from the beginning of the 50 ns simulation.** The highest fluctuations are observed for the less stable complexes due to disruption of the intermolecular contacts.

(TIF)

**S4 Fig. RMSF mapped on the 3D structures of all the CSD<sup>YB-1</sup>:oligo complexes (DNA and RNA respectively at the top and the bottom of the figure).** In red, highly flexible regions and in blue, the less flexible ones.

(TIF)

**S5 Fig. DNA nucleotides.** Root mean square deviation (RMSD) for all atoms of CSD<sup>YB-1</sup>:oligo-nucleotide along the MD trajectories starting from the beginning of the 50 ns simulation. Plots are average values of three independent run of MDS. The highest fluctuations are observed for the less stable complexes.

(TIF)

**S6 Fig. RNA nucleotides.** Root mean square deviation (RMSD) for all atoms of CSD<sup>YB-1</sup>:oligo-nucleotide along the MD trajectories starting from the beginning of the 50 ns simulation. Plots are average values of three independent run of MDS. The highest fluctuations are observed for the less stable complexes.

(TIF)

**S7 Fig. Electrophoretic analysis of YB-1 and CSD used in EMSA.** Before and after 15 minute incubation at 30°C under EMSA conditions and subjected to 13% SDS-PAGE in a tris-tricine buffer system and stained with Coomassie brilliant blue of 1 µg of YB-1 (lane 1 and 2 respectively) and 0.5 µg of CSD (lane 3 and 4 respectively).

(TIF)

**S1 Table. H-bonds network between different nucleotides and amino acids at the YB-1 CSD surface.**

(DOCX)

**S2 Table. MM/PBSA calculations results (the solute dielectric constant is equal to 4).**

(DOCX)

**S3 Table. Interactions with participation of OH groups of ribose.**

(DOCX)

## Acknowledgments

We thank the French Embassy in Moscow and the Groupe de Recherche International (GDRI) “Early Events in Human Diseases (EEHD)” for strong support of the collaboration between the INSERM UMR 829 (France) and The Institute for Protein Research (Russia).

## Author Contributions

Conceived and designed the experiments: VK PM. Performed the experiments: VK DL. Analyzed the data: VK SN LO DL PM. Contributed reagents/materials/analysis tools: VK SN LO DL PC PM. Wrote the paper: VK NV PM. Provided lab facilities to VK and PM: PC. Provided connection between VK, SN, LO and PM: PC.

## References

1. Matsumoto K, Bay BH (2005) Significance of the Y-box proteins in human cancers. *J Mol Genet Med* 1: 11–17. PMID: [19565008](#)
2. Skabkin MA, Liabin DN, Ovchinnikov LP (2006) [Nonspecific and specific interaction of Y-box binding protein 1 (YB-1) with mRNA and posttranscriptional regulation of protein synthesis in animal cells]. *Mol Biol (Mosk)* 40: 620–633.
3. Eliseeva IA, Kim ER, Guryanov SG, Ovchinnikov LP, Lyabin DN (2011) Y-box-binding protein 1 (YB-1) and its functions. *Biochemistry (Mosc)* 76: 1402–1433.
4. Hasegawa SL, Doetsch PW, Hamilton KK, Martin AM, Okenquist SA, Lenz J, et al. (1991) DNA binding properties of YB-1 and dbpA: binding to double-stranded, single-stranded, and abasic site containing DNAs. *Nucleic Acids Res* 19: 4915–4920. PMID: [1923758](#)
5. Gaudreault I, Guay D, Lebel M (2004) YB-1 promotes strand separation in vitro of duplex DNA containing either mispaired bases or cisplatin modifications, exhibits endonucleolytic activities and binds several DNA repair proteins. *Nucleic Acids Res* 32: 316–327. PMID: [14718551](#)
6. Skabkin MA, Evdokimova V, Thomas AA, Ovchinnikov LP (2001) The major messenger ribonucleoprotein particle protein p50 (YB-1) promotes nucleic acid strand annealing. *J Biol Chem* 276: 44841–44847. PMID: [11585833](#)
7. Evdokimova VM, Wei CL, Sitikov AS, Simonenko PN, Lazarev OA, Vasilenka KS, et al. (1995) The major protein of messenger ribonucleoprotein particles in somatic cells is a member of the Y-box binding transcription factor family. *J Biol Chem* 270: 3186–3192. PMID: [7852402](#)

8. Das S, Chattopadhyay R, Bhakat KK, Boldogh I, Kohno K, Prasad R, et al. (2007) Stimulation of NEIL2-mediated oxidized base excision repair via YB-1 interaction during oxidative stress. *J Biol Chem* 282: 28474–28484. PMID: [17686777](#)
9. Shibahara K, Uchiumi T, Fukuda T, Kura S, Tominaga Y, Maehara Y, et al. (2004) Targeted disruption of one allele of the Y-box binding protein-1 (YB-1) gene in mouse embryonic stem cells and increased sensitivity to cisplatin and mitomycin C. *Cancer Sci* 95: 348–353. PMID: [15072594](#)
10. Sommerville J, Baird J, Turner BM (1993) Histone H4 acetylation and transcription in amphibian chromatin. *J Cell Biol* 120: 277–290. PMID: [8421048](#)
11. Soop T, Nashchekin D, Zhao J, Sun X, Alzhanova-Ericsson AT, Björkroth B, et al. (2003) A p50-like Y-box protein with a putative translational role becomes associated with pre-mRNA concomitant with transcription. *J Cell Sci* 116: 1493–1503. PMID: [12640034](#)
12. Chansky HA, Hu M, Hickstein DD, Yang L (2001) Oncogenic TLS/ERG and EWS/Fli-1 fusion proteins inhibit RNA splicing mediated by YB-1 protein. *Cancer Res* 61: 3586–3590. PMID: [11325824](#)
13. Raffetseder U, Frye B, Rauen T, Jurchott K, Royer HD, Jansen PL, et al. (2003) Splicing factor SRp30c interaction with Y-box protein-1 confers nuclear YB-1 shuttling and alternative splice site selection. *J Biol Chem* 278: 18241–18248. PMID: [12604611](#)
14. Minich WB, Maidebura IP, Ovchinnikov LP (1993) Purification and characterization of the major 50-kDa repressor protein from cytoplasmic mRNP of rabbit reticulocytes. *Eur J Biochem* 212: 633–638. PMID: [8462540](#)
15. Davydova EK, Evdokimova VM, Ovchinnikov LP, Hershey JW (1997) Overexpression in COS cells of p50, the major core protein associated with mRNA, results in translation inhibition. *Nucleic Acids Res* 25: 2911–2916. PMID: [9207042](#)
16. Evdokimova VM, Kovrigina EA, Nashchekin DV, Davydova EK, Hershey JW, Ovchinnikov LP. (1998) The major core protein of messenger ribonucleoprotein particles (p50) promotes initiation of protein biosynthesis in vitro. *J Biol Chem* 273: 3574–3581. PMID: [9452484](#)
17. Nekrasov MP, Ivshina MP, Chernov KG, Kovrigina EA, Evdokimova VM, Thomas AA, et al. (2003) The mRNA-binding protein YB-1 (p50) prevents association of the eukaryotic initiation factor eIF4G with mRNA and inhibits protein synthesis at the initiation stage. *J Biol Chem* 278: 13936–13943. PMID: [12582179](#)
18. Evdokimova V, Ruzanov P, Imataka H, Raught B, Svitkin Y, Ovchinnikov LP, et al. (2001) The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer. *EMBO J* 20: 5491–5502. PMID: [11574481](#)
19. Svitkin YV, Evdokimova VM, Brasey A, Pestova TV, Fantus D, Yanagiya A, et al. (2009) General RNA-binding proteins have a function in poly(A)-binding protein-dependent translation. *EMBO J* 28: 58–68. doi: [10.1038/emboj.2008.259](#) PMID: [19078965](#)
20. Sutherland BW, Kucab J, Wu J, Lee C, Cheang MC, Yorida E, et al. (2005) Akt phosphorylates the Y-box binding protein 1 at Ser102 located in the cold shock domain and affects the anchorage-independent growth of breast cancer cells. *Oncogene* 24: 4281–4292. PMID: [15806160](#)
21. Evdokimova V, Ruzanov P, Anglesio MS, Sorokin AV, Ovchinnikov LP, et al. (2006) Akt-mediated YB-1 phosphorylation activates translation of silent mRNA species. *Mol Cell Biol* 26: 277–292. PMID: [16354698](#)
22. Kloks CP, Spronk CA, Lasonder E, Hoffmann A, Vuister GW, Grzesiek S, et al. (2002) The solution structure and DNA-binding properties of the cold-shock domain of the human Y-box protein YB-1. *J Mol Biol* 316: 317–326. PMID: [11851341](#)
23. Theobald DL, Mitton-Fry RM, Wuttke DS (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* 32: 115–133. PMID: [12598368](#)
24. Yamanaka K, Fang L, Inouye M (1998) The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation. *Mol Microbiol* 27: 247–255. PMID: [9484881](#)
25. Zasedateleva OA, Krylov AS, Prokopenko DV, Skabkin MA, Ovchinnikov LP, Kolchinsky A, et al. (2002) Specificity of mammalian Y-box binding protein p50 in interaction with ss and ds DNA analyzed with generic oligonucleotide microchip. *J Mol Biol* 324: 73–87. PMID: [12421560](#)
26. Max KE, Zeeb M, Bienert R, Balbach J, Heinemann U (2007) Common mode of DNA binding to cold shock domains. Crystal structure of hexathymidine bound to the domain-swapped form of a major cold shock protein from *Bacillus caldolyticus*. *FEBS J* 274: 1265–1279. PMID: [17266726](#)
27. Max KE, Zeeb M, Bienert R, Balbach J, Heinemann U (2006) T-rich DNA single strands bind to a pre-formed site on the bacterial cold shock protein Bs-CspB. *J Mol Biol* 360: 702–714. PMID: [16780871](#)
28. Sachs R, Max KE, Heinemann U, Balbach J (2012) RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution. *RNA* 18: 65–76. doi: [10.1261/ma.02809212](#) PMID: [22128343](#)

29. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* 4: 435–447.
30. Mackerell AD Jr., Feig M, Brooks CL 3rd (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25: 1400–1415. PMID: [15185334](#)
31. MacKerell AD, Bashford D, Bellott, Dunbrack RL, Evanseck JD, Field MJ, et al. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B* 102: 3586–3616. doi: [10.1021/jp973084f](#) PMID: [24889800](#)
32. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79: 926–935.
33. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126: 014101. PMID: [17212484](#)
34. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81: 3684–3690.
35. Doreleijers JF, Vranken WF, Schulte C, Lin J, Wedell JR, Penkett CJ, et al. (2009) The NMR restraints grid at BMRB for 5,266 protein and nucleic acid PDB entries. *J Biomol NMR* 45: 389–396. doi: [10.1007/s10858-009-9378-z](#) PMID: [19809795](#)
36. Shen Y, Bax A (2007) Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J Biomol NMR* 38: 289–302. PMID: [17610132](#)
37. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26: 283–291.
38. Lim VI, Kliashtorny VG (2006) [Kinetic, energetic and stereochemical factors determining molecular recognition of proteins and nucleic acids]. *Mol Biol (Mosk)* 40: 572–579.
39. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* 51: 69–82. doi: [10.1021/ci100275a](#) PMID: [21117705](#)
40. Brooks BR, Brooks CL 3rd, Mackerell AD Jr., Nilsson L, Petrella RJ, Roux B, et al. (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30: 1545–1614. doi: [10.1002/jcc.21287](#) PMID: [19444816](#)
41. Sitkoff D, Sharp KA, Honig B (1994) Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *The Journal of Physical Chemistry* 98: 1978–1988.
42. Guryanov SG, Filimonov VV, Timchenko AA, Melnik BS, Kihara H, Kutysenko VP, et al. (2013) The major mRNP protein YB-1: structural and association properties in solution. *Biochim Biophys Acta* 1834: 559–567. doi: [10.1016/j.bbapap.2012.11.007](#) PMID: [23220387](#)
43. Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image Processing with ImageJ. *Biophotonics International* 11: 36–42.
44. Dunn OJ (1959) Estimation of the Medians for Dependent Variables. *The Annals of Mathematical Statistics* 30: 192–197.
45. Cohen J (1988) *Statistical power analysis for the behavioral sciences* (2nd ed.). In: Hillsdale NJ, Lawrence Earlbaum Associates Press. 567 p.
46. Bouvet P, Matsumoto K, Wolffe AP (1995) Sequence-specific RNA recognition by the Xenopus Y-box proteins. An essential role for the cold shock domain. *J Biol Chem* 270: 28297–28303. PMID: [7499328](#)