BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Molecular Determinants for Unphosphorylated STAT3 Dimerization Determined by Integrative Modeling

Jacopo Sgrignani,[†] Simon Olsson,[†,∥] Dariusz Ekonomiuk,[†] Davide Genini,[‡] Rolf Krause,[⊥] Carlo V. Catapano,[‡] and Andrea Cavalli*[,†,§]

[†]Institute of Research in Biomedicine (IRB) and Universitá della Svizzera italiana (USI), Via Vincenzo Vela 6, CH-6500 Bellinzona, Switzerland

[‡]IOR Institute of Oncology Research, Via Vincenzo Vela 6, CH-6500 Bellinzona, Switzerland

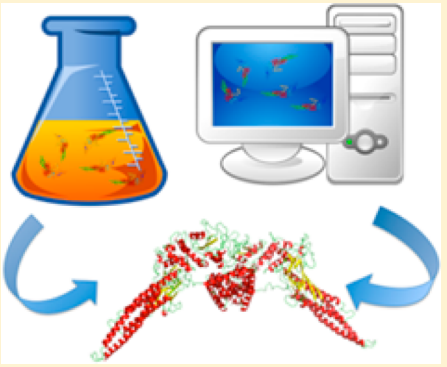
[§]Department of Chemistry, University of Cambridge, Lansfield Road, Cambridge CB2 1EW, U.K.

[∥]Laboratorium für Physikalische Chemie, Eidgenössische Technische Hochschule Zürich, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland

[⊥]Institute of Computational Science, Faculty of Informatics, Universitá della Svizzera Italiana (USI), Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

Supporting Information

ABSTRACT: Signal transducer and activator of transcription factors (STATs) are proteins that can translocate into the nucleus, bind DNA, and activate gene transcription. STAT proteins play a crucial role in cell proliferation, apoptosis, and differentiation. The prevalent view is that STAT proteins are able to form dimers and bind DNA only upon phosphorylation of specific tyrosine residues in the transactivation domain. However, this paradigm has been questioned recently by the observation of dimers of unphosphorylated STATs (USTATs) by X-ray, Förster resonance energy transfer, and site-directed mutagenesis. A more complex picture of the dimerization process and of the role of the dimers is, thus, emerging. Here we present an integrated modeling study of STAT3, a member of the STAT family of utmost importance in cancer development and therapy, in which we combine available experimental data with several computational methodologies such as homology modeling, protein−protein docking, and molecular dynamics to build reliable atomistic models of USTAT3 dimers. The models generated with the integrative approach presented here were then validated by performing computational alanine scanning for all the residues in the protein−protein interface. These results confirmed the experimental observation of the importance of some of these residues (in particular Leu78 and Asp19) in the USTAT3 dimerization process. Given the growing importance of USTAT3 dimers in several cellular pathways, our models provide an important tool for studying the effects of pathological mutations at the molecular and/or atomistic level, and in the rational design of new inhibitors of dimerization.



Signal transducer and activator of transcription factors (STATs) make up a large family of proteins (STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b, and STAT6) with lengths ranging from 750 to 850 amino acids that are key in a large number of cellular processes such as growth, survival, and differentiation.[1−3] Among all the members of the STAT family, STAT3 plays a pivotal role in carcinogenesis and cellular development by regulating the transcription of genes involved in proliferation, differentiation, apoptosis, angiogenesis, and immuno-inflammatory processes.[4] Accordingly, there is great interest in the development of selective STAT3 inhibitors, and several organic molecules that can modulate STAT3 activity,[5] frequently showing an anticancer effect, have been sought and identified.[6−9]

As in the case of other members of the STAT family, STAT3 has been only partially investigated from the structural point of view. In fact, only few and incomplete X-ray structures are currently available in the Protein Data Bank (PDB).[10−13] These studies[10−12] show, however, that STAT3 shares the same structural elements as the other members of the STAT family and is formed by five domains: (1) N-terminal domain (N-term), (2) coiled-coil domain, (3) DNA-binding domain, (4) SH2 domain, and (5) transactivation domain (TAD) (Figure 1).

Phosphorylation of tyrosine residues located in the SH2 and TAD domains of STAT proteins by tyrosine kinases (PTKs) has been identified as the critical event in the modulation of the protein function.[3,14] In the case of STAT3, phosphorylation of tyrosine 705 (Tyr705) by receptor-associated (such as JAK) and non-receptor-associated (such as Src) PTKs is the key

Received: December 17, 2014
Revised: August 1, 2015
Published: August 18, 2015

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

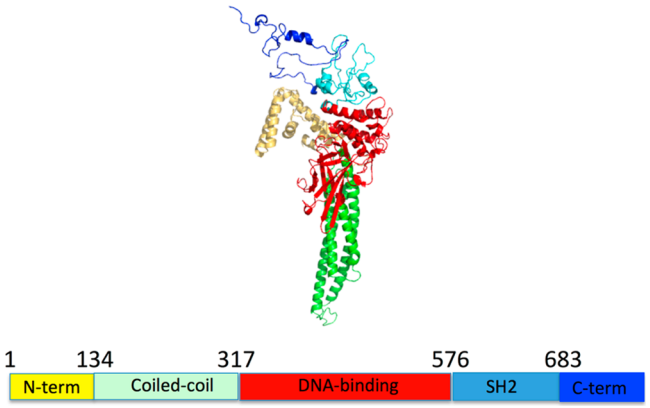


Figure 1. Full-length STAT3 general structure scheme. Protein domains are represented with different colors: N-terminal domain (yellow), coiled-coil domain (green), DNA-binding domain (red), SH2 domain (cyan), and transactivation domain (blue). Notably, this model has been built by homology modeling only for explicative purposes.

event that triggers dimerization, nuclear translocation, and the activation of its transcriptional functions.[2]

This led, over the years, to the oversimplified view that STAT3 is be able to form dimers, migrate into the nucleus, and activate transcription only after Tyr705 phosphorylation and that Tyr705-phosphorylated dimers represent the only functional form of this protein. A novel more complex picture is, however, emerging, and this paradigm has been questioned.

In fact, STAT3 can undergo several other post-translational modifications (PTM), besides Tyr705 phosphorylation, that are also key for its function under physiological and pathological conditions. These PTM include phosphorylation of serine 727 by serine protein kinases as well as lysine acetylation and methylation by protein acetyltransferase and methyltransferases.[2,15] Furthermore, the presence of unphosphorylated STAT proteins (USTATs, i.e., proteins that are not phosphorylated at key residues such as Tyr705 in the case of STAT3) inside the nucleus as well as their ability to dimerize and bind DNA has been demonstrated experimentally.[12,16−19] Notably, while their exact role in transcriptional activation and their ability to bind DNA are still being debated,[12,16] USTATs and, in particular, USTAT3 have been shown to play an important role in chromatin organization and topology, microtubule stability, and mitochondrial metabolic pathways.[2,15,20]

Recently, Wilderspin and co-workers reported the results of a multidisciplinary study that aimed to investigate the three-dimensional structure of USTAT3 dimers in complex with DNA.[12] In this study, they showed that USTAT3 dimers were able to bind DNA fragments like phosphorylated STAT3. It should, however, be noted that the N-terminal domain was removed and replaced with a green fluorescent protein and is, therefore, missing in their final model.

More recently, Müller-Newen and co-workers used Förster resonance energy transfer (FRET) and single-point mutagenesis to elucidate the molecular assembly of full-length USTAT3 dimers not interacting with DNA.[21] Interestingly, their findings support a direct interaction between the two N-terminal domains, in agreement with previous studies reporting the significance of the N-terminal domain for the dimerization of USTATs.[22]

Given the increased level of interest in the dimerization of USTAT3 and the availability of novel biophysical data, in this study we used an integrative modeling approach to investigate the structure and dynamics of this macromolecular assembly.[23,24] In particular, using various computational techniques, such as homology modeling, molecular dynamics, and protein−protein docking, together with all the available experimental information,[10,21,25,26] we were able to build and investigate reliable atomistic models of the full-length USTAT3 dimer.

METHODS

STAT3β Model. Two different STAT3 isoforms (α and β) arising from different kinds of splicing have been characterized. Their main structural difference is in the length of the C-terminal domain (~50 residues in STAT3α and ~7 residues in STAT3β).[27]

In this study, we focused our computational investigations on the STAT3β isoform that lacks the TAD domain. This because the TAD domain is, probably, not directly involved in dimerization or polymerization of the majority of STATs;[3,28,29] it is intrinsically disordered and the least conserved among all STAT subtypes. Moreover, it has been shown that STAT3β interacts with DNA and forms dimers like the full-length protein.[12]

The sequence of human STAT3 was obtained from the UniProt database (accession number P40763) and truncated at residue 722 to fit the STAT3β sequence.[27] A template search and sequence alignment (Figure 2) were performed using the HHPRED server (http://toolkit.tuebingen.mpg.de/hhpred).[30]

HHPRED is a server for remote homology detection and structure prediction based on pairwise comparison of profile Hidden Markov Models (HMMs).[31] In few words, HHPRED detects viable templates for homology modeling by comparing the HMM of the query sequence with a precompiled database of HMMs computed for all proteins with the known structure deposited in the PDB.

Two templates where selected on the basis of the determined structural elements and the resolution of the coordinates deposited in the PDB, namely, the structure of USTAT1 (PDB entry 1YVL, chain A, resolution of 3 Å)[26] and the structure of the STAT3β homodimer bound to DNA (PDB entry 1BG1, chain A, resolution of 2.25 Å).[10] These two templates where then used to generate 100 homology models with the MODELLER9.14 software package.[32] These models were ranked according the internal scoring function of the program, and finally, the quality of the best model was assessed using the PSVS, ProsaWeb, and SwissModel servers.[33−36]

Molecular Dynamics (MD) Simulations. The best-ranked model from the MODELER calculations was used as the starting point for several MD simulations. The simulations were performed using the ff14SB force field available in Amber14.[37,38] The overall protein charge was neutralized by adding four Na^+ ions using the parameters proposed by Cheatham and co-workers.[39] Finally, the protein was solvated in a box of TIP3P[40] water molecules. The size of the water box was chosen to have a minimal distance of 8 Å from the protein surface.

van der Waals and short-range electrostatic interactions were calculated within a cutoff of 8 Å, whereas the long-range electrostatic forces were taken into account with the particle mesh Ewald method.[41] Pressure and temperature were kept fixed using the Monte Carlo barostat and the Berendsen thermostat, respectively.[42] During the calculations, all bonds

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

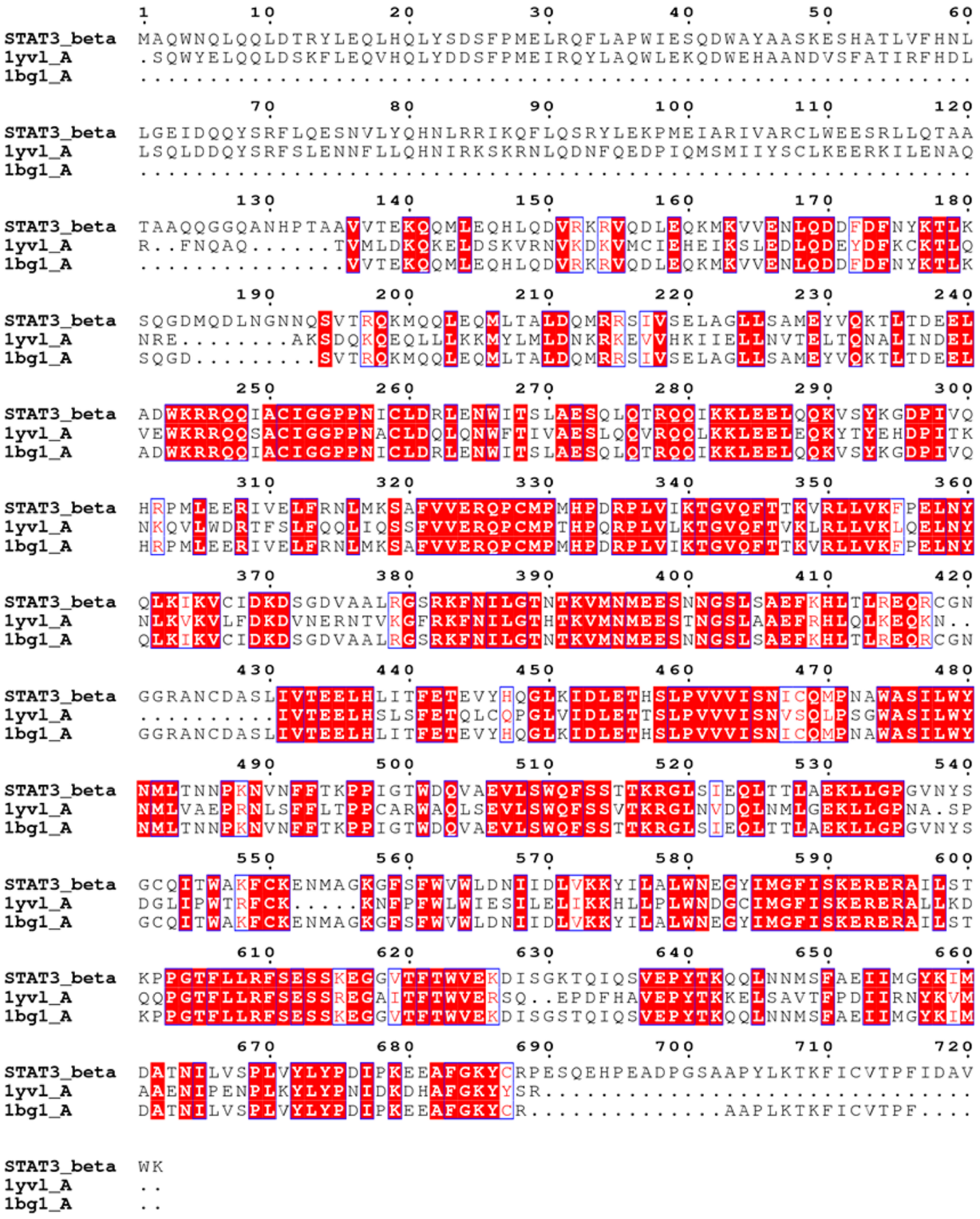


Figure 2. Sequence alignment between STAT3B and the two templates considered in this study, 1YVL and 1BG1, corresponding to three-dimensional structures of STAT1 and STAT3, respectively.

Table 1. Summary of the MD Simulations Performed during This Study

acronym	description	simulation time
USTAT3B_MD	simulation of USTAT3 homology modes	200 ns and 4 × 100 ns for testing monomer stability
STAT1MD	simulation of STAT1 structure (PDB entry 1YVL)	100 ns
BZD_MD	simulation of best dimer model selected by Zdock score	350 ns and 5 × 100 ns for an independent run for testing dimer stability
BZR_MD	simulation of best dimer model selected by Zrank score	350 ns and 5 × 100 ns for an independent run for testing dimer stability

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

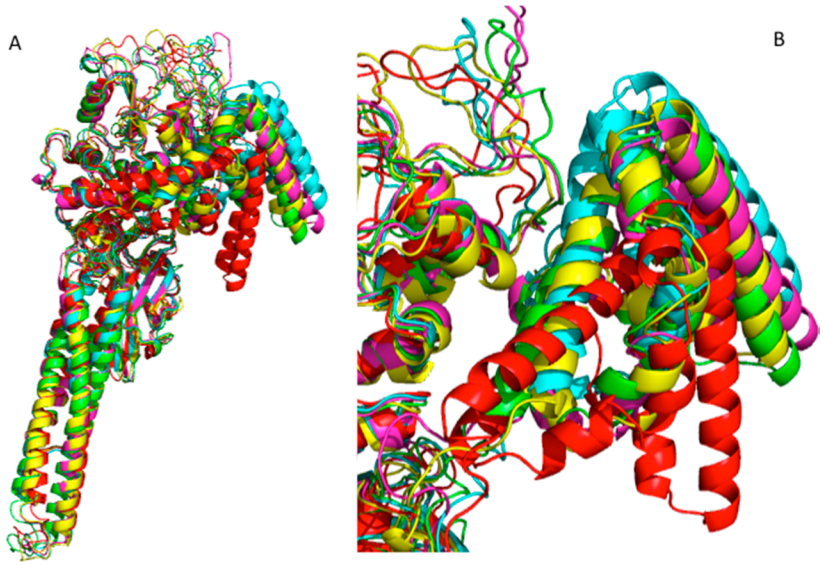


Figure 3. (A) Representative structures from the four more populated clusters (yellow, cyan, green, and magenta). The starting conformation is colored red. (B) Close-up of the N-terminal domain region.

involving hydrogen atoms were constrained by applying the SHAKE[43] algorithm. All the calculations were run with the PMEMD code in the GPU accelerated version using a time step of 2 fs.[44] An identical simulation protocol was used for the simulation of STAT1, starting from PDB structure 1YVL,[26] and for the USTAT3 dimer models (Table 1).

To remove possible steric clashes present in the initial structure, before starting MD, the entire system was first minimized for 1000 steps with a harmonic restraint on backbone atom positions (k = 20 kcal mol^{-1} Å^{-2}) and, then, for 5000 steps with all the atoms free to move. After this minimization, the system was equilibrated for 50 ps, with backbone atom coordinates restrained in the same way as in the first minimization. During equilibration, the temperature was slowly increased from 0 to 298.5 K and the pressure ramped to the reference value of 1 atm. Once the system was equilibrated, all the restraints were removed and a production run of 200 ns was performed. After this production run, a series of four shorter simulations of 100 ns each were conducted, starting from last snapshot of the long trajectory, to assess the stability and further relax the obtained structure.

Protein−Protein Docking. Protein−protein docking calculations were conducted using the Zdock program.[45] In a preliminary analysis of MD simulations of the USTAT3β monomer, we noticed that, while the general fold of the protein was stable, the orientation of the N-terminal domain as well as the conformation of the C-terminal region was quite variable. In the Zdock procedure, the proteins are treated as a rigid body disregarding, therefore, the possibility of a mutual rearrangement upon binding. Thus, to take into account this limitation, the docking procedure was repeated, in analogy to other computational studies,[24,46−49] several times starting from representative structures extracted from the MD trajectory. These representative structures were obtained by performing a cluster analysis of all conformations sampled during the last 150 ns of the MD trajectory. Clustering was conducted using the average-linkage method implemented in the ptraj program and available in the AmberTools 14 suite.[50]

The average-linkage algorithm is an iterative, agglomerative clustering method. At the beginning of the calculation, each protein conformation is placed in its own singleton cluster. Then the number of clusters is reduced through a series of iterations in which the most similar clusters are merged. The cluster similarity is assessed by computing the mean distance between members of each cluster, where the distance between structures is given by the root-mean-square deviation (rmsd) of the atomic positions after optimal superposition. This process continues until the number of clusters reaches a predetermined value (in this case 10).

The centers of the four most populated clusters were used as input in Zdock calculations.

From experimental data reported in the literature,[21,25] we could identify four essential features of the USTAT3 dimers. (i) N-Terminal domains are in direct contact. (ii) Three residues, namely, Met28, Val77, and Leu78, play an important role in this homotropic N-terminal interaction. (iii) The C-terminal parts of the two monomers are close in the space. (iv) The SH2 domain plays an important role in the strengthening of the protein−protein affinity.

The 8000 structures obtained from Zdock were, therefore, filtered to exclude all those in which Met28, Val77, and Leu78 were not part of the dimer interface.

The 102 remaining structures where then ranked according to the Zdock or Zrank scores.[45,51,52] Finally, the structure with the best Zdock score (BZD) and the one with the best Zrank score (BZR) were selected and further investigated by MD. For these two models, we first performed a long MD simulation of 350 ns using the same setup that was used for the monomer (Table 1). Then the stability of the final conformation was then further assessed by performing five independent 100 ns MD runs for each dimer model, reassigning the velocities with a different seed to the last MD snapshot.

The ZDOCK scoring function,[51] used to rank the final models, is a knowledge-based statistical potential (i.e., energy functions obtained from the analysis of a database of known structures of protein complexes), while the ZRANK[52] score is a linear combination of several energetic contributions, including electrostatics, van der Waals, and solvation.

Additional Analysis. Computational alanine scanning was performed using the DrugScorePPI server (http://cpclab.uni-

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

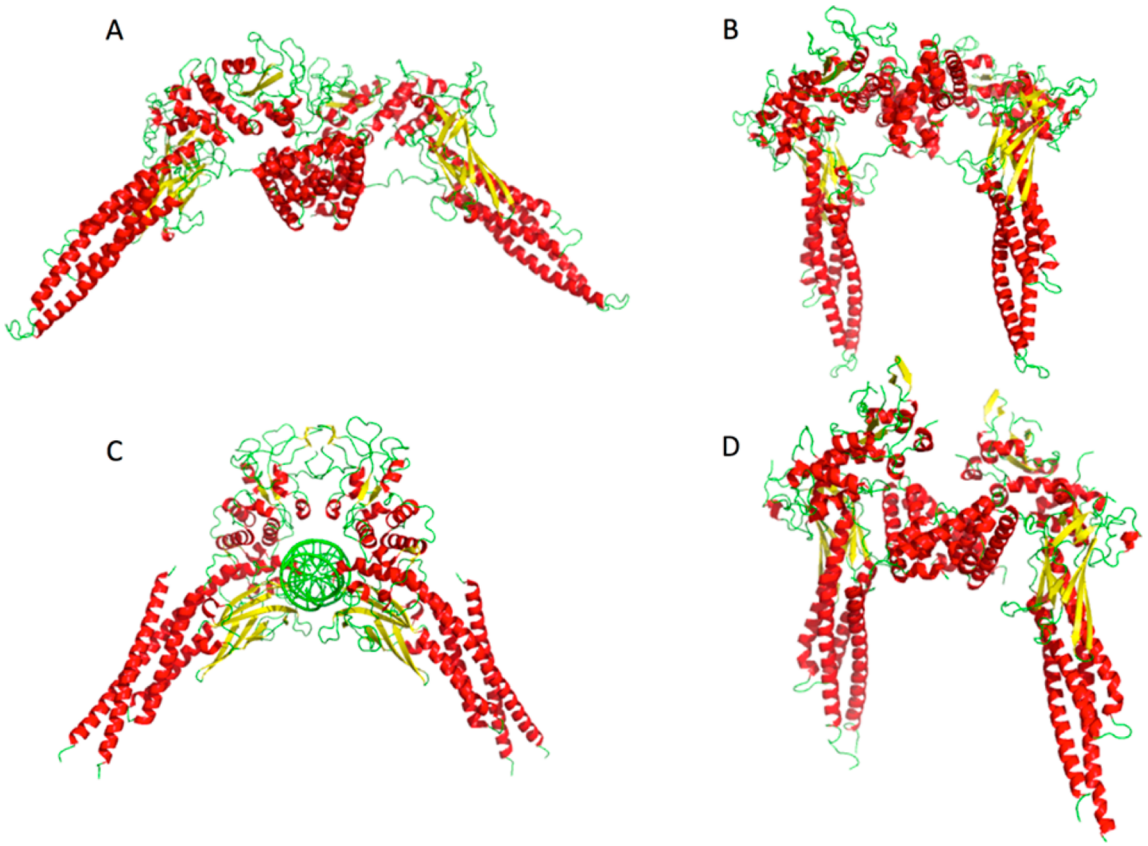


Figure 4. STAT3β homodimer models resulting from Zdock calculations: (A) BZD model, (B) BZR model, (C) USTAT3β dimer in complex with DNA (PDB entry 4E68), and (D) USTAT1 homodimer (PDB entry 1YVL).

duesseldorf.de/dsppi/).[53] The DrugScorePPI server uses a knowledge-based scoring function to compute the Gibbs free energy difference ($\Delta\Delta G$) resulting from systematic single-point mutations to alanine of all the residues that belong to the protein−protein interface. Residues were included in the interface if at least one atom was within 5 Å of the other protein in the complex.

To take into account the conformational fluctuations of the complex, the results were averaged over several representative cluster centers (to account for >50% of the sampled structure) selected, with the average-linkage method described before, from the structures sampled during an MD simulation starting from BZD and BZR, respectively.

As suggested by Gholke and co-workers,[54,55] the results of DrugScorePPI calculations were compared and integrated with those obtained from the MM-GBSA effective binding energy decomposition (i.e., the difference of the sum of the gas phase and solvation free energy calculated for the two isolated monomers and the dimer) performed with the MMPBSA.py module of Amber14 (additional details about these calculations are available as Supporting Information).[56] This computational procedure is similar to computational alanine scanning, where the residues are effectively mutated, and can be used estimate the contribution of single residues to the total dimerization energy.[54,57]

RESULTS AND DISCUSSION

Two templates were identified by the HHPRED server (Figure 2) as being suitable for modeling STAT3β, namely the STAT1 and STAT3β (from *Mus musculus*) structures from PDB entries 1YVL[26] and 1BG1,[10] respectively. The STAT1 structure has a coverage of 95% (i.e., 95% of the protein residues have a correspondence in the template), and the sequence identity with STAT3β is ∼53%, a value well within the range required for reliable homology modeling.[58] On the other hand, the sequence of 1BG1 is identical, but it covers only 80% of the residues of full-length STAT3β.

The 100 models generated by MODELER, using the two templates simultaneously, were structurally compared to investigate the variability among them. The structural comparison was conducted by first aligning the backbone atoms of all the models to the average structure and then by computing the α-carbon rmsd and the angle among the centers of mass of the N-terminal, DNA-binding, and SH2 domains.

The rmsd and interdomain angle (reported in Figure S1), as well as visual inspection, confirmed that all the models were very similar with minor differences only in the conformation of the long loop connecting the N-terminal domains with the DNA-binding domains, and in the structure of the intrinsically disordered C-terminus.

Accordingly, the best model was selected on the basis of the internal scoring function of MODELER, and its quality was checked with some specific tools (see Methods).

The *Z*-score obtained from ProsaWeb analysis (−11.76) and the recently proposed QMEAN *Z*-score[59] (0.68) were both within the range expected for X-ray structures of similar sizes (Figures S2 and S3). Furthermore, the model showed a very low percentage (0.8%) of residues in disallowed regions of the Ramachandran plot, with violations mainly located in loops or disordered regions of the C-terminal part of the SH2 domain and of the transactivation domain (Figure S4).

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

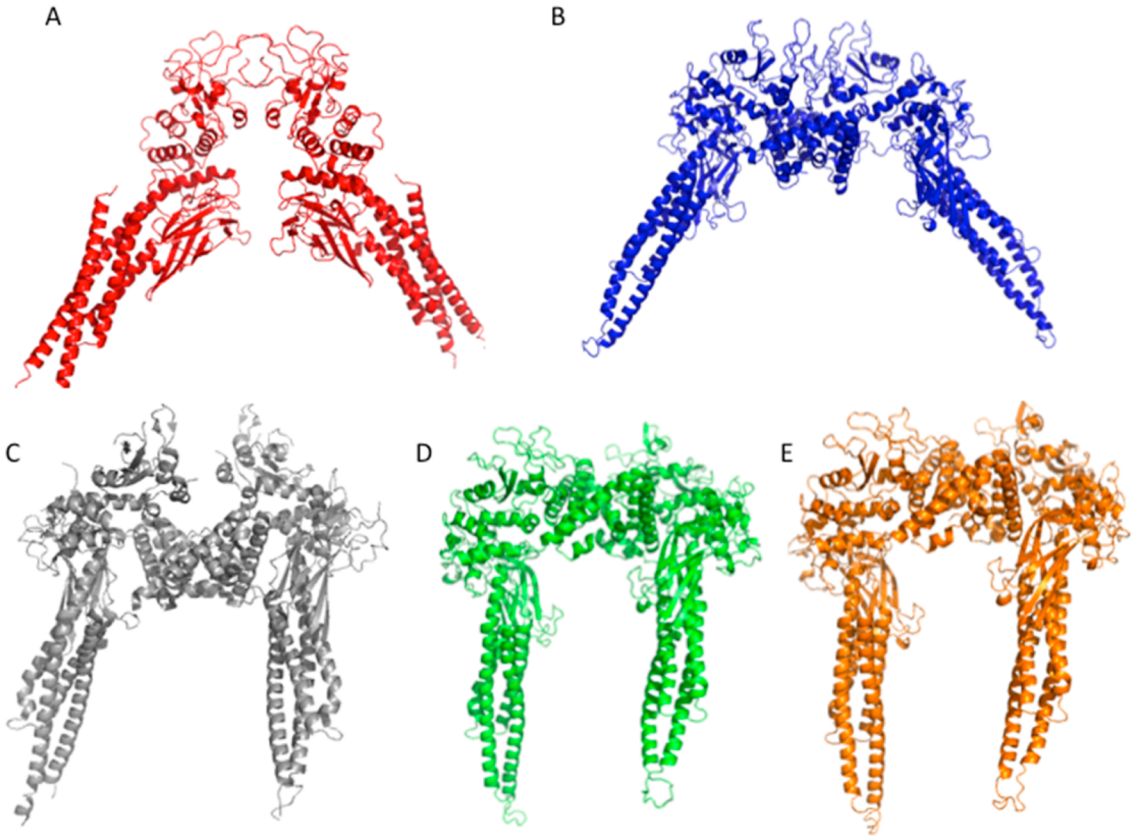


Figure 5. BZD representative structures from the more populated cluster (A) and USTAT3β X-ray dimer structure (PDB entry 4E68) (B; in this case, the DNA molecule has been omitted for the sake of clarity). USTAT1 homodimer (C, PDB entry 1YVL) and BZR representative structure from the first two more populated clusters (D and E).

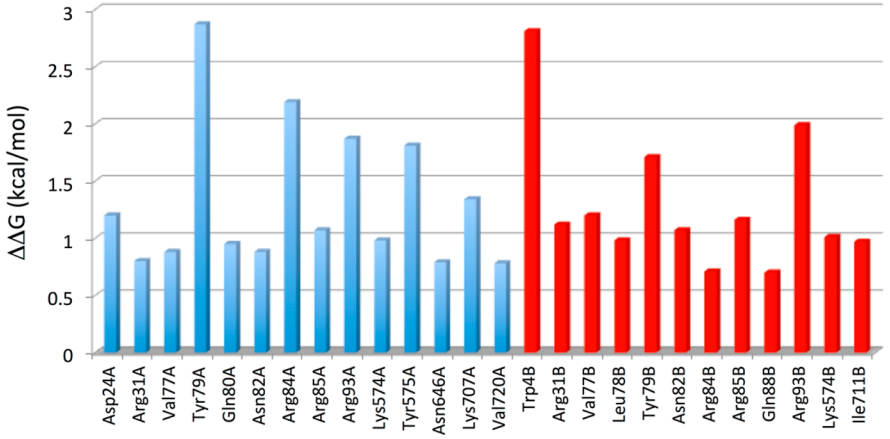


Figure 6. $\Delta\Delta G$ values calculated with the DrugScorePPI server for all the residues forming the BZD dimer interface considering the representative structure from the more populated cluster. All the residues with $\Delta\Delta G$ values of >0.7 kcal/mol were reported. Positive $\Delta\Delta G$ values indicate potential hot spot residues. Values from chain A are colored blue, while those from chain B are colored red. Chains A and B represent two identical monomers.

To study its dynamics and sample enough conformations for protein−protein docking calculations, the selected model was then simulated for 200 ns in explicit water.

After 50 ns, the system reached an equilibrium state. In fact, the average rsmd value, with respect to the mean structure calculated over the last 150 ns of the simulation, was 2.48 (0.4) Å, while with respect to the initial structure, it was 6.46 (1.0) Å (Figure S5A−C).

These values suggested, however, some degree of conformational variability of the STAT3β structure. Therefore, we tested the stability of the equilibrated structure by running four additional independent MD simulations of 100 ns each. The results of these simulations were compared with those from the initial 200 ns MD trajectory, taking into account both the rmsd with respect to the starting model coordinates and the angle among the N-terminal, DNA-binding, and SH2 domains. This analysis confirmed the stability of the conformation obtained at the end of the 200 ns simulation. In fact, the configuration space explored during these four additional simulations was smaller than that sampled in the first run (Figures S5−S7).

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

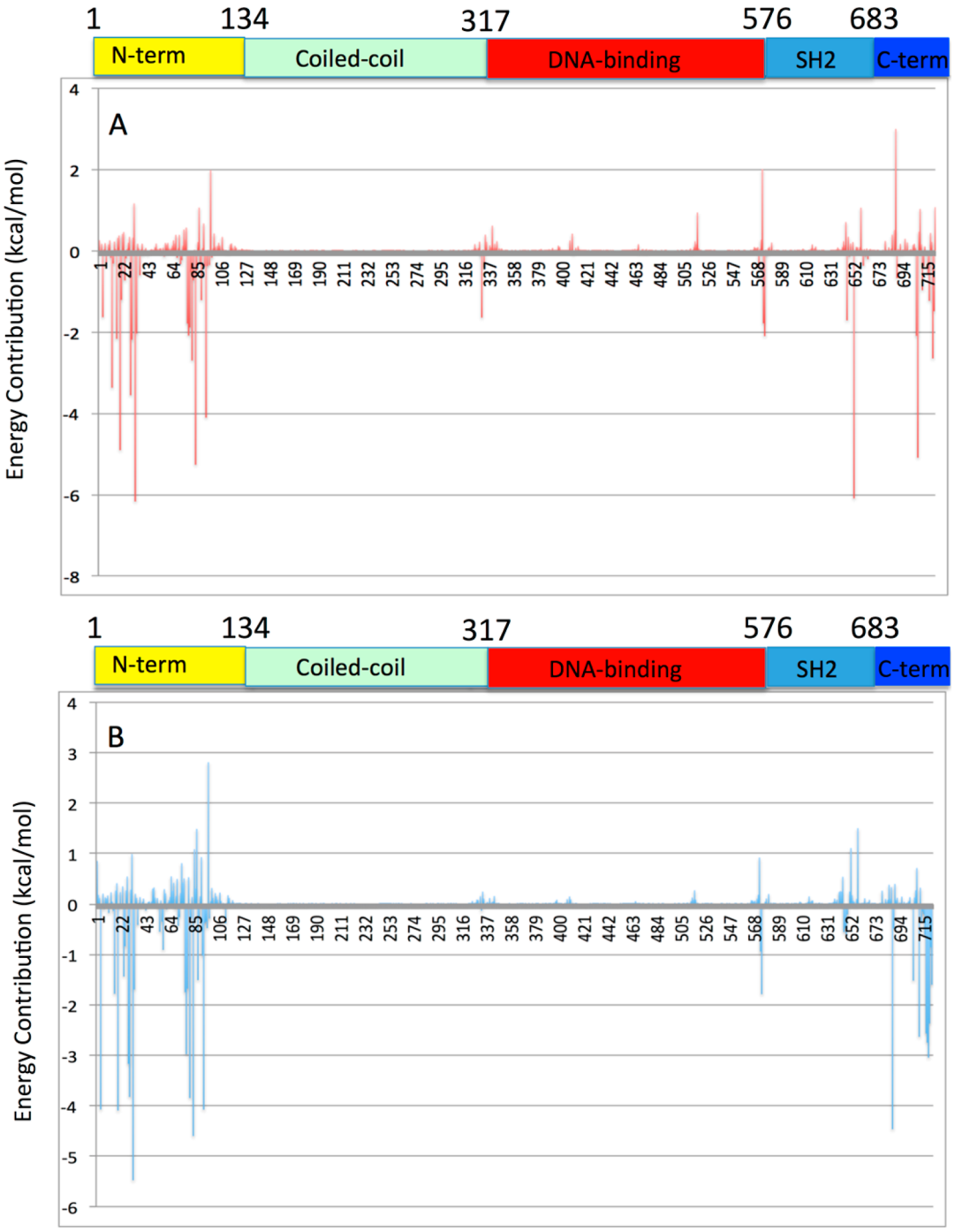


Figure 7. Contribution to the total dimerization energy given by every residue in chain A (A) and chain B (B) of the BZD model. For the sake of clarity, a focused view of the two graphs is reported (C and D); in this case, only residues having a contribution of less than −1 kcal/mol are shown. The standard error in the mean is <0.1 kcal/mol for every residue.

To gain further insight into the dynamics of the protein, we performed a cluster analysis of the conformations collected in the last 150 ns of the 200 ns production run.

From this analysis, four highly populated clusters were identified with a total of >80% of the sampled conformations (Figure 3). Visual comparison of these structures revealed that the N-terminal domain has a slightly different orientation with respect to that assumed in the starting model but interacts with the same protein domains (Figure S8B). Notably, this conformation was maintained over all the four 100 ns MD simulations that were conducted to check both the stability of the model and the convergence of our calculations.

For the sake of comparison, we also performed a 100 ns MD simulation of STAT1 (PDB entry 1YVL), which is the only crystal structure in which the N-terminal domain is determined together the other protein domains. This simulation was conducted manly to determine the time scale required to sample different N-terminal domain orientations. Interestingly, the STAT1 simulation converges to a structure in which the N-terminal (Figure S8A) domain has an orientation that is different from the STAT3β most populated clusters, as well as

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

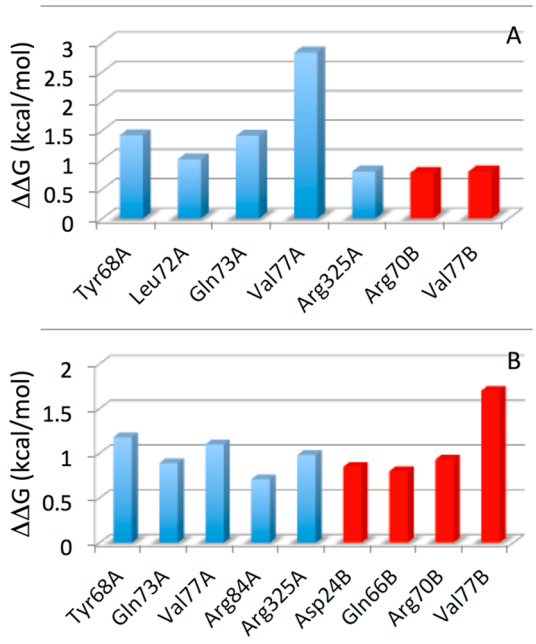


Figure 8. $\Delta\Delta G$ values calculated with the DrugScorePPI server for residues as the BZR dimer interface considering the representative structure from the two more populated clusters (A and B). All the residues with $\Delta\Delta G$ values of >0.7 kcal/mol were reported. Values from chain A are colored blue while those from chain B red. Chains A and B represent two identical monomers.

the initial STAT1 X-ray structure. This observation suggests that the motions of the STAT3 N-terminal domain can be sampled in simulations of similar lengths.

At this point, to assess if using more models, among those generated by homology modeling, would lead to a better exploration of the STAT3 conformational space, we compared the structure of the 100 models with those of 10 representative structures sampled during MD and selected by cluster analysis (Table S1). Moreover, we computed the per-residue root-mean-square fluctuation (rmsf) for both the 100 models and the snapshots sampled during the last 150 ns of the MD simulation (Figure S9).

The first analysis showed that all the models have similar Cα rmsd values, if compared with the same representative MD structures (standard deviation of <0.5 Å). The per-residue rmsf plot (Figure S9) showed, moreover, larger fluctuations for the MD ensemble of structures than in the homology models.

In summary, these analyses indicated that (1) all the models are very similar when compared with the representative structure from MD and cover a similar conformational space; (2) MD simulations sample a conformational space larger than that explored by all 100 homology models.

Docking Calculations of STAT3β Homodimers. Despite the growing interest in USTAT3 dimers driven by the recent observation that even in its unphosphorylated form they play a key role in several cellular functions, a complete atomistic picture is still lacking.[12,16−19] To address this problem, we modeled USTAT3 dimers following an approach that combines experimental observations available in the literature[12,21] with rigid body protein−protein docking and MD simulations.

Recent FRET experiments conducted by Müller-Newen and co-workers[21] suggested that full-length USTAT3 molecules form homodimers in which the N-terminal domains of the two monomers are directly in contact and the C-terminal domains are close in space. From these observations, they proposed a model with a parallel orientation of the two monomers in the USTAT3 dimers. To improve the understanding of the molecular determinants of USTAT3 dimerization, they also investigated the effects of two point mutations: L78R and R609Q. The L78R mutant was found to be unable to form dimers, thus supporting the significance of the N-terminal domain for USTAT3 dimerization. In the case of R609Q, the emerging picture was less clear. However, the acquired data suggested a significant role of the SH2 domain in the modulation of the protein−protein affinity.[21]

Given the importance of the N-terminal domain for USTAT3 dimerization, we searched the literature for potential experimental evidence suggesting interchain interaction of N-terminal domains.

A complete STAT3 structure is still lacking. However, Kuriyan and co-workers studied the N-terminal homotropic interactions in STAT4, suggesting a possible interaction surface on the basis of X-ray and mutagenesis studies.[25] In particular, from these investigations, two other residues, Met28 and Phe77 (this residue is replaced with valine, a similar hydrophobic residue, in STAT3), emerged as being important for the dimerization. Notably, Chen and co-workers found a similar interaction surface and binding orientation for the two N-terminal domains of STAT1.[26]

On the basis of these findings, we filtered the 8000 (2000 for every representative conformation from the four most populated clusters) structures generated by Zdock to discard those in which the identified residues (i.e., Met28, Val77, and Leu78) were not in the dimer interface, yielding 102 possible solutions. Finally, from these structures, two dimer models were selected on the basis of the best Zdock score (BZD; score of 1351) and best Zrank score (BZR; score of −79) (Figure 4).

The structures of two selected models differ considerably both in the protein−protein interface region and in the orientation of the coiled-coil domains. The two models were selected by rescoring the same conformational ensemble with two different energy functions. The structural differences are, therefore, a consequence of slight differences in the way structural features are taken into account in the Zdock and Zrank scores, respectively. To assess the reliability of the two models we used, therefore, a very long MD simulation of 350 ns was conducted to quantify their stability. The structures were, moreover, validated with the help of experimental data available from the literature.

Comparison of the Complex Models with Existing Experimental Data. All the STAT dimer structures deposited in the PDB are incomplete or of other STAT subtypes.

The similarity analysis was conducted for the DNA-bound USTAT3β dimer (PDB entry 4E68),[12] with the homotropic STAT4 N-terminal dimer proposed by Kuriyan and co-workers[25] and with one of the biological assemblies proposed by Mao et al.[26] for STAT1 (PDB entry 1YVL).

The global arrangement of the BZD model is very similar to that of the USTAT3β dimer in complex with DNA (PDB entry 4E68), with only two minor differences. First, the two N-terminal domains are located in the place occupied, in the crystal structure, by the DNA fragment (Figure 4). Second, the two coiled-coil domains have the same orientation but are farther from each other than in all the other known dimer structures. On the other hand, the overall shape of the BZR model is more similar to that of the USTAT1 dimer (PDB entry 1YVL).

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

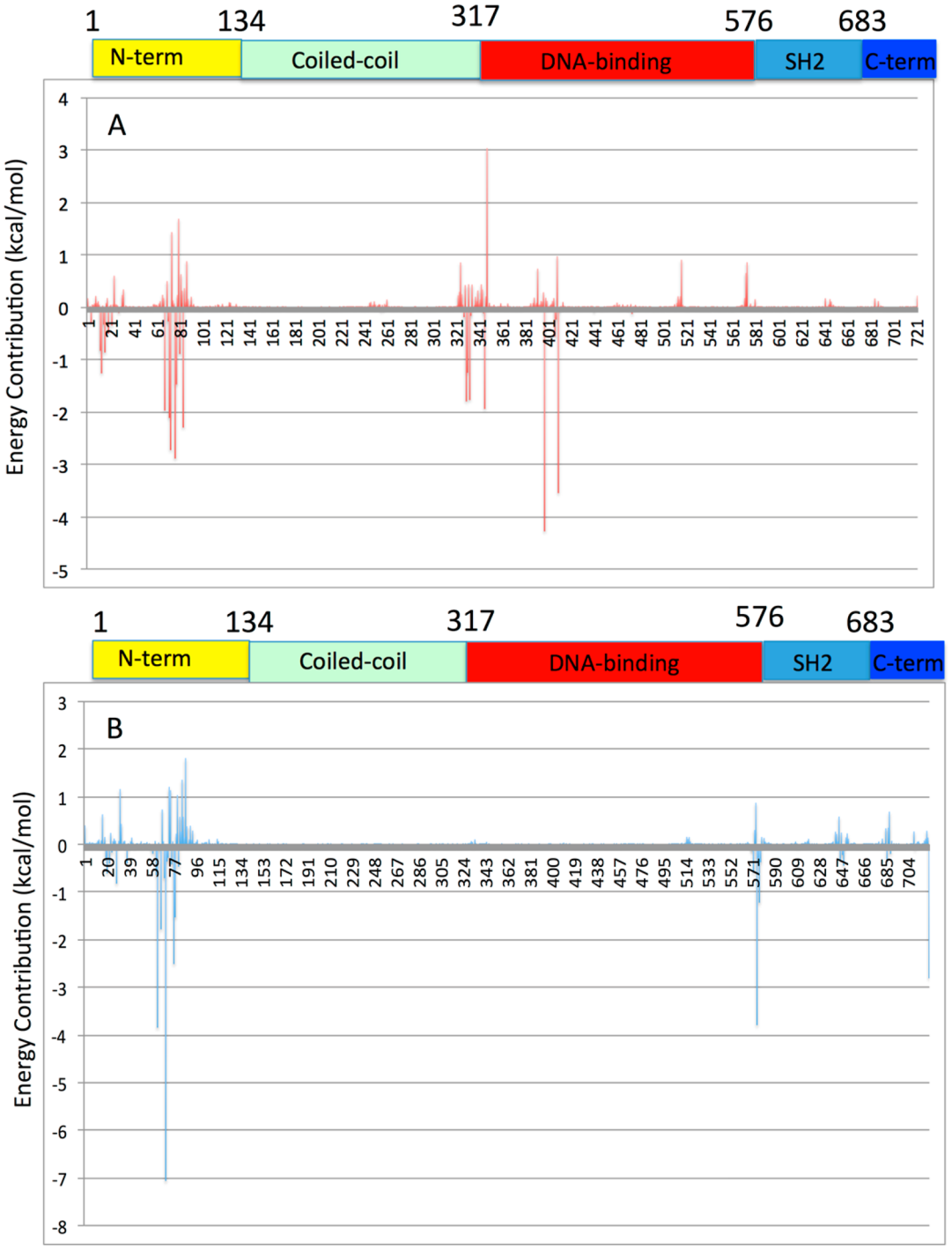


Figure 9. Contribution to the total dimerization energy given by every residue in chain A (A) and chain B (B) of the BZR model. For the sake of clarity, a focused view of the two graphs is reported (C and D); in this case, only residues having a contribution of less than −1 kcal/mol are shown. The standard error in the mean is <0.1 kcal/mol for every residue.

The relative orientation of the N-terminal domains in the BZD model fits well both models of biological assembly obtained from the structure of USTAT1[26] and STAT4[25] [Cα rmsds of 4.2 and 4.4 Å, respectively (Figure S10)]. In the case of the BZR model, while the overall shape is similar, the fit is poorer [rmsd of 7.7 Å for both proposed structures (Figure S10)].

Molecular Dynamics Simulation of Dimers. In the Zdock modeling approach, each protein is treated as a rigid body, precluding, therefore, the possibility of structural rearrangements upon docking. To mitigate the effects of this approximation, several independent docking runs were conducted with representative conformations of USTAT3 sampled during the MD simulation and selected by cluster analysis. However, while this strategy can be helpful to better explore the dimer conformational space, this does not contribute to the relaxation of the final dimers. Thus, to partially overcome this limitation and to take into account the possible mutual induced-fit effects of the single monomers, we simulated both BZD and BZR for 350 ns.

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

Visual inspection and rmsd calculation (Figure S11) for both the entire structure and the dimerization surface showed that, during the simulation, the structure of the BZD model did not significantly change with respect to the its starting geometry. In particular, the N-terminal domains were found to be very stable in their initial position, while the coiled-coil domains were more mobile. In fact, the distance between these two domains significantly decreased during the simulation, going from ∼200 Å to a final value of ∼160 Å, leading to a structure closer to the DNA-bound USTAT3β[12] dimer than to the BZD starting model (Figure 5A,B).

Although relatively long, a 350 ns simulation might be too short for a complete relaxation of a macromolecular object of the size of a STAT3 dimer. Therefore, to assess the stability or metastability of the conformations adopted by the dimer at the end of the 350 ns simulation, we performed 10 independent runs of 100 ns each (five for BZD and five for BZR). All simulations were started from the last snapshot of the 350 ns long MD simulation with different initial velocities.

Finally, we evaluated the Cα rmsd of the STAT3 region more involved in the dimerization process (i.e., N-terminal and SH2 domains). The results of this analysis confirmed the stability of the BZD model and showed a lower stability for the BZR models (Figure S12).

We performed a cluster analysis of the MD trajectories to identify representative snapshots. In the case of the BZD model, the results of this analysis indicated that the most populated cluster covers 54% of the entire sampled conformational space, while for the BZR model, there are two highly populated clusters with 27 and 26% of the structures (Figure 5).

Dimerization Energy Decomposition. Several residues of the N-terminal domain were identified, in mutagenesis studies, as being key for USTAT3 dimer formation. In particular, the L78R mutation was reported to impair USTAT3 dimerization.[21] Moreover, given the high degree of homology between STAT3 and STAT4 and the work of Kuriyan and co-workers,[25] it appears that mutation of Met28 or Val77 could reduce the affinity between the two N-terminal domains.

Assessing the effect of several single-point mutations on protein dimerization using long time scale MD simulations is still a prohibitive task, despite recent developments in computational algorithms and more powerful hardware. However, less time-consuming computational methodologies that use knowledge-based or approximated energetic calculations[53,56] have been developed, thus making it possible to predict the change in the binding free energy of a protein−protein complex with a reasonable level of accuracy in a few minutes or hours.[53,60]

In this study, we used the DrugScorePPI empirical function and MM-GBSA per-residue energy decomposition to evaluate the energetic contribution of residues located in the protein−protein interface. The results of these predictions were averaged over several snapshots selected by cluster analysis to take into account the dimer dynamics.

In the case of the BZD model (Figure 6), DrugScorePPI calculations showed that the protein−protein interface is formed by ∼60 residues from the N-terminal and SH2 domains. Moreover, Arg31, Val77, Tyr79, Asn82, Arg84, and Arg85, of both chains A and B, contribute significantly to the dimerization energy.

In addition to that, MM-GBSA energy decomposition analysis (Figure 7 and Table S2) indicated that two protein regions of the N-terminal domains (one from Met1 to Gln32 and the other from Ser75 to Glu100) play an important role in the strengthening of the protein−protein interaction.

Comparing the results of our calculation with those of mutagenesis investigations reported in the literature that identified Met28 and Leu78[21,25] as possible hot spots, we noted that these two residues are included in the two regions discussed before.

It should be noted that during the selection of the dimer models, the presence of these two residues in the dimerization interface was used as a filter. While this is certainly a bias, it does not, *a priori*, imply that Met28 and Leu78 make a large contribution to the dimerization energy.

In addition to that, we searched the scientific literature for additional experimental data to further validate our findings. Interestingly, we found an accurate investigation of the STAT4 dimerization process conducted by Ota et al.[61] This work confirms the strong propensity of STAT proteins to form dimers by homotropic interaction between N-terminal domains. Moreover, in addition to Leu78, they found that Asp19 is important for protein dimerization (i.e., its mutation to Arg completely suppresses dimer formation). In a manner consistent with these data, the histidine at position 19 in STAT3 makes a very large contribution to the dimerization energy during MM-GBSA analysis.

Finally, the region from Lys685 to Phe710 across the SH2 and C-terminal domains is known to play an important role in the stabilization of the dimer. Interestingly, a residue, Tyr705, that according to the MM-GBSA analysis significantly increases the STAT3 dimerization propensity is also within this region of the protein.

For the BZR model (Figure 8), the protein−protein interface was found to be smaller and comprised only ∼40−50 residues. In this case, from calculation with both MM-GBSA and DrugScorePPI, it appears that the initial region of the protein (from Met1 to Gln32), which contains the experimentally identified hot spot Met28 and His19, does not significantly contribute to the dimerization energy. In fact, the residues with the largest contribution to the dimerization energy are clustered in a region delimited by Arg70 and Leu90. Moreover, in this case, the MMBGSA energy profile (Figure 9 and Table S3) was more asymmetric (i.e., the two monomers show a different contribution) than that obtained for the BZD domain and the contribution coming from the SH2 and N-terminal domains is almost zero.

SUMMARY AND CONCLUSIONS

USTAT3 dimers recently emerged as important biological adducts for STAT3-related pathways.[12,16,21] In this study, we applied an integrative modeling approach that uses available experimental data[10,21,26] together with computational methods to build (i) a USTAT3β model and (ii) to select possible USTAT3β homodimer models (BZD and BZR).

Molecular dynamics simulations were conducted to relax the models and explore their stability and dynamics. Furthermore, computational alanine scanning and MM-GBSA energy decomposition were used to assess the role of individual interface residues in the stabilization of USTAT3β dimers.

These calculations highlighted a larger dimerization interface for the BZD model. For both models (BZD and BZR), a region, from Arg70 to Leu90, with potential dimerization hot

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

spots was identified by performing computational alanine scanning. These findings are in agreement with the null dimerization propensity observed for mutants L78R of STAT3 and F77A of STAT4, reported by Muller-Newen[21] and Kuryan,[25] respectively. Furthermore, while this work was under revision, a new structure of N-terminal STAT3 dimers was published and made available in the PDB (entry 4ZIA).[62] Thus, we compared the crystal structure of the N-terminal dimers with the BZD and BZR models. In both cases, we found that the N-terminal domains have an orientation that is almost identical to that in the X-ray structure, with Cα rmsd values of 4.5 and 4.8 Å for BZD and BZR, respectively (Figure S13).

However, only for the BZD model was a significant contribution to the dimerization energy estimated for the initial part of the N-terminal domain that includes His19 and Met28 (other experimentally identified hot spots[25]) and the SH2 domain, which is also known to play an important role in strengthening the protein−protein affinity.[21]

Both proposed models are compatible with the experimental information about the shape of USTAT3β dimers from the studies of Müller-Newen and co-workers.[21] However, given (i) the large contribution to the dimerization energy of the residue in the SH2 domain, (ii) the high degree of similarity of the BZD N-terminal domain−N-terminal domain dimer with that proposed by Kuriyan and co-workers,[25] (iii) the better agreement with single-point mutagenesis data, and (iv) the higher stability showed during the MD simulations, the BZD model emerges as a better candidate of a representative structure of USTAT3 dimers.

The structures described in this study are, to the best of our knowledge, the first attempt to obtain a structural model, at an atomic level of detail, of a full USTAT3β dimer not interacting with DNA. Additional experimental and theoretical work is required to further validate and refine these models. However, although approximate, they should provide a valuable tool for the interpretation of past experiments and the rational design of future biochemical investigations. In fact, these models could be used in structure-based drug design projects aimed to identify new pharmacological tools to modulate USTAT3 dimers and explore this strategy as a novel route for therapeutic invention.

In the future, the models described here should, therefore, contribute to a better understanding of the cellular function of USTAT3 dimers, the mechanism of action of small molecule STAT3 inhibitors, and the effect of pathological STAT3 mutations that are being increasingly found in cancer and other diseases.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/bi501529x.

Additional analyses (rmsd and N-terminal, DNA binding, and SH2 domain angle) for homology models and simulation outputs; results form *Z*-score analysis; localization of the residue outside permitted zones of the Ramachandran plot; pictures of more populated STAT1 clusters; N-terminal domain alignment with respect to data from refs 30 and 62; initial and final conformations of BZR model; and additional details about energy decomposition calculations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Institute for Research in Biomedicine, Via Vincenzo Vela 6, CH-6500 Bellinzona, Switzerland. E-mail: andrea.cavalli@irb.usi.ch.

Funding

This work was supported by "Fondazione per la Ricerca e la Cura dei Linfomi in Ticino". We acknowledge the Swiss National Supercomputing Center (CSCS) for the availability of high-performance computing resources. S.O. is funded by an Independent Postdoc grant from The Danish Council for Independent Research|Natural Sciences (DFF-4002-00151).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. Rajiv Chopra and Dr. David A. Frank for kindly making available the STAT3 N-terminal domain X-ray structure before the official release in the PDB.

■ ABBREVIATIONS

MD, molecular dynamics; STAT3, signal transducer and activator of transcription factor, subtype 3; USTAT3, unphosphorylated STAT3; BZD, best dimer model selected by the Zdock score; BZR, best dimer model selected by the Zrank score; TAD, transactivation domain.

■ REFERENCES

(1) Levy, D. E., and Darnell, J. E., Jr. (2002) Stats: transcriptional control and biological impact. *Nat. Rev. Mol. Cell Biol. 3*, 651−662.

(2) Yu, H., Pardoll, D., and Jove, R. (2009) STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat. Rev. Cancer 9*, 798−809.

(3) Lim, C. P., and Cao, X. (2006) Structure, function, and regulation of STAT proteins. *Mol. BioSyst. 2*, 536−550.

(4) Xiong, A., Yang, Z., Shen, Y., Zhou, J., and Shen, Q. (2014) Transcription Factor STAT3 as a Novel Molecular Target for Cancer Prevention. *Cancers 6*, 926−957.

(5) Siveen, K. S., Sikka, S., Surana, R., Dai, X., Zhang, J., Kumar, A. P., Tan, B. K. H., Sethi, G., and Bishayee, A. (2014) Targeting the STAT3 signaling pathway in cancer: role of synthetic and natural inhibitors. *Biochim. Biophys. Acta, Rev. Cancer 1845*, 136−154.

(6) Miklossy, G., Hilliard, T. S., and Turkson, J. (2013) Therapeutic modulators of STAT signalling for human diseases. *Nat. Rev. Drug Discovery 12*, 611−629.

(7) Masciocchi, D., Gelain, A., Villa, S., Meneghetti, F., and Barlocco, D. (2011) Signal transducer and activator of transcription 3 (STAT3): a promising target for anticancer therapy. *Future Med. Chem. 3*, 567−597.

(8) Lavecchia, A., Di Giovanni, C., and Cerchia, C. (2014) Novel inhibitors of signal transducer and activator of transcription 3 signaling pathway: an update on the recent patent literature. *Expert Opin. Ther. Pat. 24*, 383−400.

(9) Debnath, B., Xu, S., and Neamati, N. (2012) Small molecule inhibitors of signal transducer and activator of transcription 3 (Stat3) protein. *J. Med. Chem. 55*, 6645−6668.

(10) Becker, S., Groner, B., and Müller, C. W. (1998) Three-dimensional structure of the Stat3beta homodimer bound to DNA. *Nature 394*, 145−151.

(11) Ren, Z., Mao, X., Mertens, C., Krishnaraj, R., Qin, J., Mandal, P. K., Romanowski, M. J., McMurray, J. S., and Chen, X. (2008) Crystal structure of unphosphorylated STAT3 core fragment. *Biochem. Biophys. Res. Commun. 374*, 1−5.

(12) Nkansah, E., Shah, R., Collie, G. W., Parkinson, G. N., Palmer, J., Rahman, K. M., Bui, T. T., Drake, A. F., Husby, J., Neidle, S., Zinzalla,

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

G., Thurston, D. E., and Wilderspin, A. F. (2013) Observation of unphosphorylated STAT3 core protein binding to target dsDNA by PEMSA and X-ray crystallography. *FEBS Lett. 587*, 833−839.
(13) Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, J. E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol. 112*, 535−542.
(14) Zhong, Z., Wen, Z., and Darnell, J. E., Jr. (1994) Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science 264*, 95−98.
(15) Yu, H., Lee, H., Herrmann, A., Buettner, R., and Jove, R. (2014) Revisiting STAT3 signalling in cancer: new and unexpected biological functions. *Nat. Rev. Cancer 14*, 736−746.
(16) Timofeeva, O. A., Chasovskikh, S., Lonskaya, I., Tarasova, N. I., Khavrutskii, L., Tarasov, S. G., Zhang, X., Korostyshevskiy, V. R., Cheema, A., Zhang, L., Dakshanamurthy, S., Brown, M. L., and Dritschilo, A. (2012) Mechanisms of unphosphorylated STAT3 transcription factor binding to DNA. *J. Biol. Chem. 287*, 14192−14200.
(17) Haan, S., Kortylewski, M., Behrmann, I., Müller-Esterl, W., Heinrich, P. C., and Schaper, F. (2000) Cytoplasmic STAT proteins associate prior to activation. *Biochem. J. 345*, 417−421.
(18) Kretzschmar, A. K., Dinger, M. C., Henze, C., Brocke-Heidrich, K., and Horn, F. (2004) Analysis of Stat3 (signal transducer and activator of transcription 3) dimerization by fluorescence resonance energy transfer in living cells. *Biochem. J. 377*, 289−297.
(19) Braunstein, J., Brutsaert, S., Olson, R., and Schindler, C. (2003) STATs dimerize in the absence of phosphorylation. *J. Biol. Chem. 278*, 34133−34140.
(20) Mohr, A., Chatain, N., Domoszlai, T. s., Rinis, N., Sommerauer, M., Vogt, M., and Müller-Newen, G. (2012) Dynamics and non-canonical aspects of JAK/STAT signalling. *Eur. J. Cell Biol. 91*, 524−532.
(21) Domoszlai, T. S., Martincuks, A., Fahrenkamp, D., Schmitz-Van de Leur, H., Kuster, A., and Muller-Newen, G. (2014) Consequences of the disease-related L78R mutation for dimerization and activity of STAT3. *J. Cell Sci. 127*, 1899−1910.
(22) Neculai, D., Neculai, A. M., Verrier, S., Straub, K., Klumpp, K., Pfitzner, E., and Becker, S. (2005) Structure of the unphosphorylated STAT5a dimer. *J. Biol. Chem. 280*, 40782−40787.
(23) Degiacomi, M. T., and Dal Peraro, M. (2013) Macromolecular Symmetric Assembly Prediction Using Swarm Intelligence Dynamic Modeling. *Structure 21*, 1097−1106.
(24) Russel, D., Lasker, K., Webb, B., Velazquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012) Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol. 10*, e1001244.
(25) Chen, X., Bhandari, R., Vinkemeier, U., van den Akker, F., Darnell, J. E., and Kuriyan, J. (2003) A reinterpretation of the dimerization interface of the N-terminal Domains of STATs. *Protein Sci. 12*, 361−365.
(26) Mao, X., Ren, Z., Parker, G. N., Sondermann, H., Pastorello, M. A., Wang, W., McMurray, J. S., Demeler, B., Darnell, J. E., Jr., and Chen, X. (2005) Structural bases of unphosphorylated STAT1 association and receptor binding. *Mol. Cell 17*, 761−771.
(27) Ng, I. H. W., Ng, D. C. H., Jans, D. A., and Bogoyevitch, M. A. (2012) Selective STAT3-α or -β expression reveals spliceform-specific phosphorylation kinetics, nuclear retention and distinct gene expression outcomes. *Biochem. J. 447*, 125−136.
(28) Wenta, N., Strauss, H., Meyer, S., and Vinkemeier, U. (2008) Tyrosine phosphorylation regulates the partitioning of STAT1 between different dimer conformations. *Proc. Natl. Acad. Sci. U. S. A. 105*, 9238−9243.
(29) Droescher, M., and Vinkemeier, U. (2012) Self-association of STAT Proteins from Monomers to Paracrystals. In *Jak-Stat Signaling: From Basics to Disease* (Decker, T., and Müller, M., Eds.) pp 47−63, Springer, Vienna.
(30) Soding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res. 33*, W244−248.
(31) Hildebrand, A., Remmert, M., Biegert, A., and Soding, J. (2009) Fast and accurate automatic structure prediction with HHpred. *Proteins: Struct., Funct., Genet. 77*, 128−132.
(32) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779−815.
(33) Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins: Struct., Funct., Genet. 66*, 778−795.
(34) Sippl, M. J. (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins: Struct., Funct., Genet. 17*, 355−362.
(35) Wiederstein, M., and Sippl, M. J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res. 35*, W407−410.
(36) Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins: Struct., Funct., Genet. 50*, 437−450.
(37) Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput. 11*, 3696.
(38) Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., Cheatham, T. E., III, Darden, T. A., Duke, R. E., Gohlke, H., Goetz, A. W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T. S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K. M., Paesani, F., Roe, D. R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C. L., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., and Kollman, P. A. (2014) *AMBER14*, University of California, San Francisco.
(39) Joung, I. S., and Cheatham, T. E. (2009) Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B 113*, 13279−13290.
(40) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, L. M. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(41) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577−8593.
(42) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3690.
(43) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comput. Phys. 23*, 327−341.
(44) Le Grand, S., Götz, A. W., and Walker, R. C. (2013) SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun. 184*, 374−380.
(45) Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., and Weng, Z. (2014) ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics 30*, 1771−1773.
(46) Sgrignani, J., Bonaccini, C., Grazioso, G., Chioccioli, M., Cavalli, A., and Gratteri, P. (2009) Insights into docking and scoring neuronal alpha4beta2 nicotinic receptor agonists using molecular dynamics simulations and QM/MM calculations. *J. Comput. Chem. 30*, 2443−2454.
(47) De Leo, F., Sgrignani, J., Bonifazi, D., and Magistrato, A. (2013) Structural and dynamic properties of monoclonal antibodies immobilized on CNTs: a computational study. *Chem. - Eur. J. 19*, 12281−12293.
(48) Sgrignani, J., Bon, M., Colombo, G., and Magistrato, A. (2014) Computational Approaches Elucidate the Allosteric Mechanism of

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501

Human Aromatase Inhibition: A Novel Possible Route to Small-Molecule Regulation of CYP450s Activities? *J. Chem. Inf. Model. 54*, 2856−2868.
(49) Amaro, R. E., Baron, R., and McCammon, J. A. (2008) An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput.-Aided Mol. Des. 22*, 693−705.
(50) Roe, D. R., and Cheatham, T. E. (2013) PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, ? *J. Chem. Theory Comput. 9*, 3084−3095.
(51) Mintseris, J., Pierce, B., Wiehe, K., Anderson, R., Chen, R., and Weng, Z. (2007) Integrating statistical pair potentials into protein complex prediction. *Proteins: Struct., Funct., Genet. 69*, 511−520.
(52) Pierce, B., and Weng, Z. (2007) ZRANK: reranking protein docking predictions with an optimized energy function. *Proteins: Struct., Funct., Genet. 67*, 1078−1086.
(53) Kruger, D. M., and Gohlke, H. (2010) DrugScorePPI webserver: fast and accurate in silico alanine scanning for scoring protein-protein interactions. *Nucleic Acids Res. 38*, W480−486.
(54) Ciglia, E., Vergin, J., Reimann, S., Smits, S. H., Schmitt, L., Groth, G., and Gohlke, H. (2014) Resolving hot spots in the C-terminal dimerization domain that determine the stability of the molecular chaperone Hsp90. *PLoS One 9*, e96031.
(55) Metz, A., Schanda, J., Grez, M., Wichmann, C., and Gohlke, H. (2013) From determinants of RUNX1/ETO tetramerization to small-molecule protein-protein interaction inhibitors targeting acute myeloid leukemia. *J. Chem. Inf. Model. 53*, 2197−2202.
(56) Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA.py: An Efficient Program for End-State Free Energy Calculations? *J. Chem. Theory Comput. 8*, 3314−3321.
(57) Gohlke, H., Kiel, C., and Case, D. A. (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol. 330*, 891−913.
(58) Nayeem, A., Sitkoff, D., and Krystek, S. (2006) A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models. *Protein Sci. 15*, 808−824.
(59) Benkert, P., Tosatto, S. C., and Schomburg, D. (2008) QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Struct., Funct., Genet. 71*, 261−277.
(60) Fernández-Recio, J. (2011) Prediction of protein binding sites and hot spots. *WIREs Comput. Mol. Sci. 1*, 680−698.
(61) Ota, N., Brett, T. J., Murphy, T. L., Fremont, D. H., and Murphy, K. M. (2004) N-domain-dependent nonphosphorylated STAT4 dimers required for cytokine-driven activation. *Nat. Immunol. 5*, 208−215.
(62) Hu, T., Yeh, J. E., Pinello, L., Jacob, J., Chakravarthy, S., Yuan, G. C., Chopra, R., and Frank, D. A. (2015) Impact of the N-terminal domain of STAT3 in STAT3-dependent transcriptional activity. *Mol. Cell. Biol.*, MCB.00060-15.

DOI: 10.1021/bi501529x
Biochemistry 2015, 54, 5489−5501