J Mol Model (2009) 15:257–266
DOI 10.1007/s00894-008-0382-0

ORIGINAL PAPER

Effects of tryptophan residue fluorination on streptavidin stability and biotin–streptavidin interactions via molecular dynamics simulations

Jarosław J. Panek · Thomas R. Ward · Aneta Jezierska · Marjana Novič

Received: 8 October 2008 / Accepted: 27 October 2008 / Published online: 4 December 2008

Abstract Due to its highly specific and very strong binding, the (strept)avidin–biotin system forms the basis for numerous applications in the life sciences: immunoassays, DNA detection systems, affinity chromatography, etc. Fine-tuning of the ligand binding abilities of this system might provide new technologies with relevance to nanoscale research. Here, we report our computational investigations on wild type (WT) and modified streptavidin (SAV), assessing the impact of fluorination of tryptophan residues on biotin binding ability. Complexes of biotin with four SAV protein variants (WT-SAV, 4fW-SAV, 5fW-SAV and 6fW-SAV) were studied. We found that protein stability and folding are predicted to be weakly affected by fluorination. The host protein binding pocket decreases its ability to form numerous hydrogen bonds to biotin in the case of the 4fW-SAV variant. Conversely, the 5fW-SAV mutant is predicted to have an even more stable ligand–host hydrogen bonding network than WT-SAV. Thermodynamic perturbation investigations predict a decrease in biotin binding free energy from 3.0 to 6.5 kcal/mol per tetrameric host, with the 5fW-SAV mutant being least affected. Overall, the computational findings indicate that 6fW-SAV and, especially, 5fW-SAV to be promising variants of streptavidin for potential modifiable picomolar binding of the biotin ligand family.

Keywords Biotin · Fluorotryptophan · Molecular dynamics · Streptavidin · Thermodynamical perturbation

Electronic supplementary material The online version of this article (doi:10.1007/s00894-008-0382-0) contains supplementary material, which is available to authorized users.

J. J. Panek · A. Jezierska · M. Novič
National Institute of Chemistry,
Hajdrihova 19,
1001 Ljubljana, Slovenia

T. R. Ward
Institute of Chemistry, University of Neuchâtel,
Av. de Bellevaux 51,
2009 Neuchâtel, Switzerland

T. R. Ward
Department of Chemistry, University of Basel,
Spitalstrasse 51,
4056 Basel, Switzerland

J. J. Panek (✉) · A. Jezierska
Faculty of Chemistry, University of Wrocław,
F. Joliot-Curie 14,
50-383 Wrocław, Poland
e-mail: jarek@elrond.chem.uni.wroc.pl

Introduction

The avidin–biotin complex and related systems are good examples of the biochemical importance and strength of non-covalent interactions [1]. The binding constants for the complex formation are of the order of 10^{13} M^{-1} to 10^{15} M^{-1} [2–4]. This property places the biotin–(strept)avidin interaction among the strongest non-covalent interactions known in nature [5]. Detailed structural and biophysical investigations of the biotin–(strept)avidin system have shown that this extraordinary affinity relies on multiple H-bonding interactions, coupled with a deep hydrophobic pocket provided by aromatic residues [6–8]. The origins of the exceptional strength of the biotin–(strept)avidin interaction are found in a unique combination of the shape of the ligand (i.e. hydrophilic head with three hydrogen-bond-capable centres, hydrophobic tail) and the binding pocket of the host protein. The enthalpy of complex formation is enhanced by the presence of well-defined hydrogen bonds

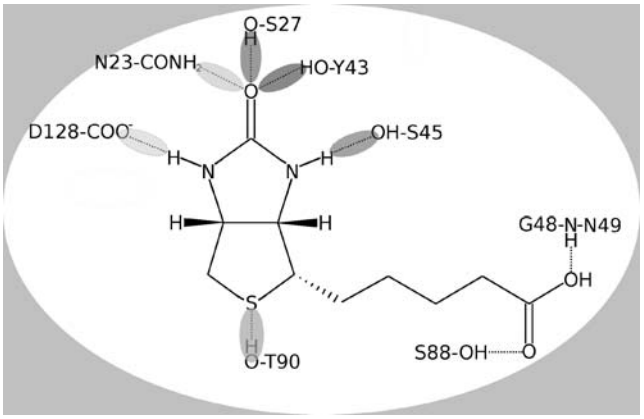

Fig. 1 Hydrogen bonding framework of the biotin–streptavidin system

as illustrated in Fig. 1. Entropic effects also play a significant role, since biotin is partially locked in place by the flexible 3–4 loop (avidin residues 35–46) [9]. The joint synergistic effect [10] of these forces leads to an exceptionally large binding affinity, well outside the classical realm of non-covalent interactions [11]. Fine-tuning of the (strept)avidin–biotin system is therefore motivated both by extending possible applications in life sciences, as well as by the urge and need to understand binding and recognition mechanisms found in nature. Efficient anchoring of native or modified biotin and analogs into the chiral environment of the protein gives rise to a whole range of potential applications: bioassays, probes, enantioselective sensors and catalysts [2, 12, 13]. Natural biotin-binding proteins belong mainly to two distinct families: (1) avidin and related, though less active, derivatives found in eggs, and (2) prokaryotic streptavidin (SAV) [14, 15]. Modifications to natural systems providing "smart" technologies [16] are motivated by future nanotechnology applications.

Generally, two possibilities of affecting (strept)avidin–biotin interactions have been distinguished:

(1) Modification of the protein;
(2) Structural changes in the ligand.

In the recent literature, two interesting computational applications of ligand modifications have been reported [17, 18]. In the first, a computational fluorine scanning study predicted a modification of the biotin skeleton able to bind strongly to avidin [17]. The second combined calculations and an experimental study on SAV. In this latter paper Dixon et al. [18] showed that 9*R*-methylbiotin, although less efficient than biotin itself, is the best among its known analogues. Following the modification schemes proposed by Dixon et al. [18], the current study reports the computational modification of SAV. The protein chosen for this purpose is a core SAV (residues 13–155), which, in its native, tetrameric form, contains four biotin-binding pockets. The pocket includes a hydrophobic cage formed primarily by four tryptophan residues. Modifications of these residues will lead to changes in the biotin-binding properties of the protein, as the hydrophobic cage has been shown to be pivotal in the enthalpy vs entropy balance of the ligand-binding process [19, 20]. The presence of this cage has been invoked as critically important for the unusually high affinity of biotin–SAV binding [21]. Stayton and coworkers [19] have presented an elegant study of single point SAV mutants in which tryptophan residues are replaced with phenylalanine (conservative mutation) or alanine. The former resulted in a 10× to 100× smaller binding constant K_a, while the latter lowered K_a by at least six orders of magnitude, thus unambiguously demonstrating the crucial importance of aromatic residues within the binding pocket. In this context, it is particularly interesting to study conservative substitutions in the hydrophobic cage. The particular modification discussed in the current study is a substitution of tryptophan with fluorinated analogues: 4-, 5- or 6-fluorotryptophan (4fW, 5fW and 6fW, respectively; see Fig. 2). Fluorination of selected residues is a methodology often used to label and/or modify protein–ligand interactions. The introduction of fluorine atoms facilitates structural investigations and ligand binding assays probed either via ^{19}F NMR [22, 23], fluorescence [24] or UV resonance Raman [25] techniques. The use of fluorotryptophan as a probe is useful only when it results in only a minor perturbation of the native system. This happens in many cases, but there are also notable exceptions, e.g. cardiac troponin-C does not change its folding when a F153 (5fW) mutation is introduced, but F104(5fW) mutant folding is dependent on the particular *Escherischia coli* strain used in the expression system [26].

The main aim of this study was to investigate the influence of tryptophan fluorination in positions 4, 5 or 6 of the amino acid on the dynamics of biotin binding in the binding pocket and on the structure of the protein itself. Hydrogen bond interactions in the binding pocket and the

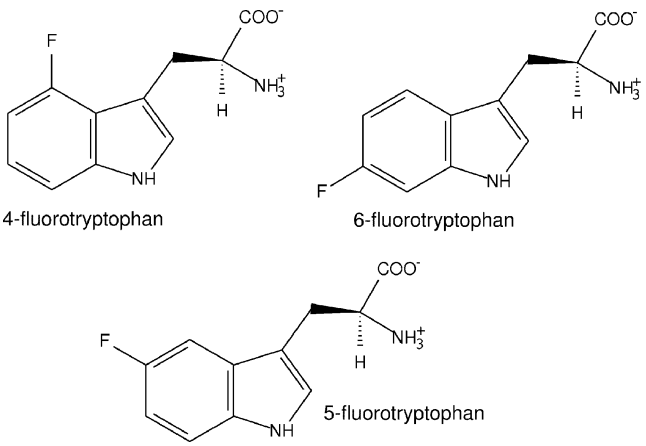

Fig. 2 Structures of 4-, 5- and 6-fluorotryptophan (4fW, 5fW, 6fW)

overall backbone structure are discussed below, as both are essential for efficient biotin binding. The electrostatic potential and solvent accessible surface area (SASA) were also analysed. Changes in protein–ligand properties through simulation time were investigated via classical molecular dynamics (MD) methods. MD simulations were performed for the wild type (WT) SAV protein as well as for its fluorinated variants. In addition, thermodynamics perturbation calculations were carried out to predict changes in the free energy of biotin binding [27].

Computational methodology

For the current study, core SAV (residues 13-155) was selected to serve as basis for both apo-SAV and the biotin–SAV complex. The preparation of models was based on available X-ray structures (at 1.58 Å resolution), solved at various pH values [28]. Initial atomic coordinates for this study were prepared using experimental data for the biotin–streptavidin complex, PDB entry: 2IZF [28]. Typical protonation states were assumed (Arg, Lys—positive, Asp, Glu—negative). Based on visual examination of the environment of the histidine residues and on pH-dependent structural studies of the SAV system [28], His87 was taken to be positive, while the His127 unit was assumed to be neutral and protonated only at the δ position. The crystal water present in the experimental PDB structure was removed. The negative net charge of the system was equalised by two sodium ions positioned in regions with minimum electrostatic potential energy. Tryptophan fluorination in positions 4, 5 or 6 was obtained by substitution of single hydrogen atoms for fluorine (see Fig. 2). This gives us a set of variants (4fW-SAV, 5fW-SAV and 6fW-SAV) in which all tryptophan residues have been substituted by their fluorinated counterparts. WT-SAV served as a reference for the set of fluorinated complexes.

To complete the set up for MD simulations, it was necessary parametrise the non-standard biotin residue. Atomic charges for the residue were parametrised according to the density functional theory (DFT) applying Becke's [29] three-parameter hybrid potential with the nonlocal correlation functional of Lee, Yang and Parr [30] denoted as B3LYP. The 6-31G(d,p) double-zeta valence-split basis set was used for geometry optimisation, and the Merz-Kollman [31] electrostatic-potential-fitting scheme was employed to assign the atomic charges for use in the molecular mechanics force field. This part of the simulation was carried out using the Gaussian03 suite of programs [32].

Subsequently, the tetrameric biotin–WT-SAV complex and its variants (biotin with 4fW-SAV, 5fW-SAV and 6fW-SAV) were placed in rectangular boxes (72×70×75 Å) filled with water. Periodic boundary conditions (PBC) were applied to simulate the bulk solution. The CHARMM22 force field [33] was employed for the organic part and counterions, while water molecules were described by the TIP3P model [34]. Non-bonded short-range electrostatic interactions were gradually switched off between 8 and 10 Å. The particle mesh Ewald method was applied to evaluate long-range electrostatic interactions [35–37]. A timestep of 2 fs was employed, and a temperature of 300 K was enforced by thermostatting. An initial minimisation was carried out for 10,000 steps to remove short contacts between the protein molecule and the solvent. Molecular dynamics runs were then performed, consisting of 0.5 ns equilibration at constant pressure and temperature controlled by Langevin thermostat and barostat [38, 39], followed by 40 ns data collection at constant volume and temperature. The lengths of bonds involving hydrogen atoms were kept fixed using the SHAKE algorithm [40]. This part of the calculations was carried out using the program NAMD [41]. Post-processing analysis of the trajectories obtained was performed using the program VMD [42]. For improved reliability of the statistical data, most analyses were carried out by averaging the data obtained from four independent biotin-binding sites of the model. Additionally, the PDB2PQR package [43] coupled with the Adaptive Poisson-Boltzman Solver [44] (APBS) was used to evaluate the electrostatic properties of WT-SAV.

Further, thermodynamic perturbation [27] investigations of the biotin-binding free energy change were performed for the fluorinated variants. The equilibrated structures previously used to initialise the MD data collection run, were reused. The computational setup was the same as described above for the MD, but the timestep was reduced to 1 fs. The transformation index λ ran from 0 (standard tryptophan) to 1 (fluorinated residue) with a step of 0.1 (and with a smaller 0.01 step in the 0–0.05 and 0.95–1 ranges). The perturbation data were collected for 0.080 ps at each λ parameter window. Both forward and backward runs were carried out for convergence testing. The NAMD program was used for the thermodynamic part of the calculations [41]. Graphical presentation of the results was obtained with the assistance of the VMD program [42].

Results and discussion

The current study investigated four complexes of biotin with WT-SAV, 4fW-SAV, 5fW-SAV and 6fW-SAV protein variants. Molecular dynamic simulations were performed to estimate the stability of protein folding and to obtain insights into the streptavidin–biotin bonding pattern.

The hydrogen bonding framework of streptavidin-bound biotin is described in Fig. 1. Apart from the bonds analysed below, there are two additional contacts formed by the

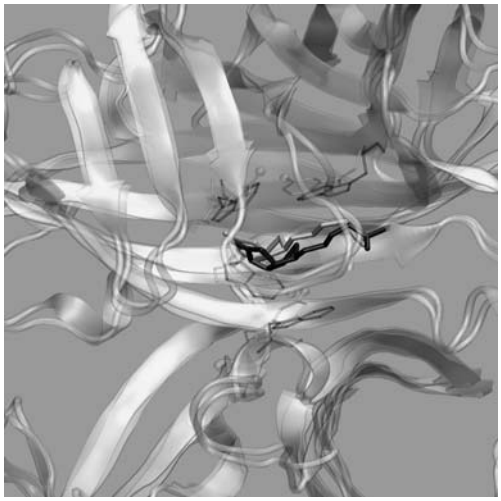

Fig. 3 Biotin in the streptavidin binding pocket. Wild type (WT) (*dark grey*) and 5-fluorotryptophan (5fW) (*light grey*) streptavidin (SAV) superimposed after 40 ns of molecular dynamics (MD) simulation. The hydrophobic cage of tryptophan residues (for 5fW, fluorine atoms are marked as *small spheres*) is shown. Additional viewing orientations are presented in Figure S1 [see electronic supplementary material (ESM)]

valeric acid carboxyl group that are not discussed further here. The hydrophobic cage of tryptophan residues, forming van der Waals interactions with biotin, is depicted in Fig. 3.

The hydrogen bonds analysed are those formed by the biotin urea moiety (HN–CO–NH) with surrounding residues. Specifically, the biotin carbonyl oxygen atom interacts with the hydroxyl groups of Ser27 and Tyr43, and the Asn23 side-chain amide group. One of the NH groups of biotin interacts with the Ser45 hydroxyl, while the other interacts with the Asp128 carboxyl group. This extensive network of hydrogen bonds has been investigated experimentally by site-directed mutation [7, 45, 46]. Surprisingly, selective disruption of a chosen interaction via introduction of a biotin ureido-carbonyl oxygen atom did not lead to large shifts in biotin position, despite large decreases in binding enthalpies. The overall structure of the protein was almost unaffected. A point mutation study of the biotin-binding pocket residue Y33 of avidin [47] yielded similar findings. This suggests that nonpolar forces play an important role in binding, as pointed out earlier [21]. Additionally, the network of hydrogen bonds is expected to exhibit cooperativity, and loss of one contact could be partially compensated by the action of remaining residues.

Any significant changes in protein secondary structure will be correlated with deviation from an initial structure based on experimental data from 2IZF. The root mean square deviation (RMSD) graphs and hydrogen bond time series for wild type (WT-SAV) and fluorinated streptavidin variants are presented in Fig. 4. As can be seen, WT-SAV and its fluorinated variants evolve remarkably close to each other in time, indicating that no dramatic changes in protein

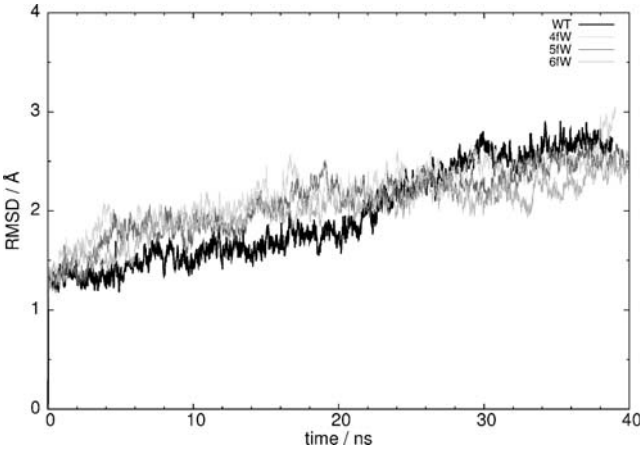

Fig. 4 The root mean square deviation (RMSD, in Å) of the protein backbone as a function of MD simulation time (in ns) for WT-SAV and fluorinated variants

structure are taking place. Moreover, the overall deviations from WT-SAV are small enough to consider the fluorinated structures to be essentially identical to WT-SAV in terms of secondary and tertiary structure (see also Fig. 3, where WT and 5fW-SAV variants are superimposed, revealing remarkable conservation of the hydrophobic cage). The length of our simulation (40 ns) is, of course, not sufficient to study the complete refolding of the structure—such a study might be well beyond the scope of current MD techniques, as timescales of microseconds or longer might be involved. In our opinion, however, the simulated time is long enough to indicate the comparative stability of the WT fold with that of fluorinated residues

Figure 3 indicates that, even after the relatively long simulation time of 40 ns, the differences between WT and the 5fW variant are not large. This observation is supported not only by the RMSD evolution discussed above, but also by another structural parameter: mass-weighted radius of gyration (RGYR, see Fig. 5). The RGYR corresponds to an

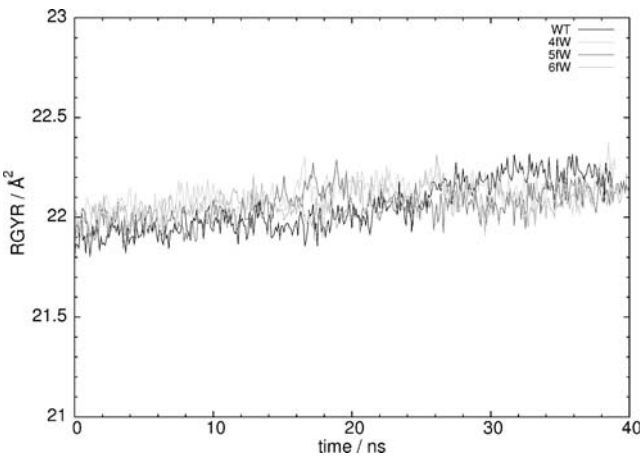

Fig. 5 Mass-weighted radius of gyration (RGYR) calculated as a function of time for WT-SAV and its variants

average moment of inertia of the protein, and its time evolution shows only small changes during the simulation. This provides additional proof that the structures of all the studied proteins do not undergo any serious changes within the timescale of the simulation. The SAV tetramer possesses quite a rigid backbone framework, which is evident in the root mean square fluctuation (RMSF) graph for individual residues (Figs. 6, 7). Figure 6 shows clearly that the binding pocket is very well conserved during simulation of the biotin–WT-SAV complex. The RMSF values do not exceed 1 Å in the inner part of the protein (see also Fig. 7, where positions of residues in contact with biotin are indicated). The C- and N-terminal parts of the chains, as well as the loops, exhibit large motions, and they are responsible for most of the observed RMSD and RMSF variability. Figure 7 shows that the difference between WT and the fluorinated variants is negligible, which is especially visible in the difference graph (where WT RMSF was subtracted from the RMSF for the variants shown). Illustrated graphically in Fig. 6, the rigidity of the binding pocket is also shown in Fig. 7. Additionally, we observe that the differences due to fluorination influence only slightly the backbone motions of the residues forming the hydrogen bonding network with biotin. In the next paragraph we examine this network in more detail.

The hydrogen bonding network formed around the biotin ureido ring consists of three contacts formed by the carbonyl oxygen and one contact for each of the imino groups (see Fig. 1). The former interactions link the oxygen atom to the hydroxyl groups of Ser27 and Tyr43 as well as the amide group of Asn23. The latter are N–H···O bonds formed between biotin nitrogen atoms and the Ser45 hydroxyl oxygen or Asp128 carboxyl group. Additionally, a contact between the sulphur atom of the biotin and the hydroxyl group of Thr90 also stabilises the position of the ligand. Statistical analysis of the time evolution of these contacts is provided in Table 1 and is discussed below. We prefer not to probe in detail the presence of possible weak C–H···O interactions, which are considered to be important for protein folding [48, 49] and in specific cases might also be involved in ligand binding [50]. An extensive experimental study [51] has shown C–H···O bonds to be in general ubiquitous, but rather restricted to ligand–water interactions in protein complexes. According to the latter study, ligands prefer to use stronger hydrogen bond capabilities for binding to protein residues, leaving the weaker functionalities to interact with water. Accordingly, the C–H···O bonds are beyond the scope of this study; moreover,

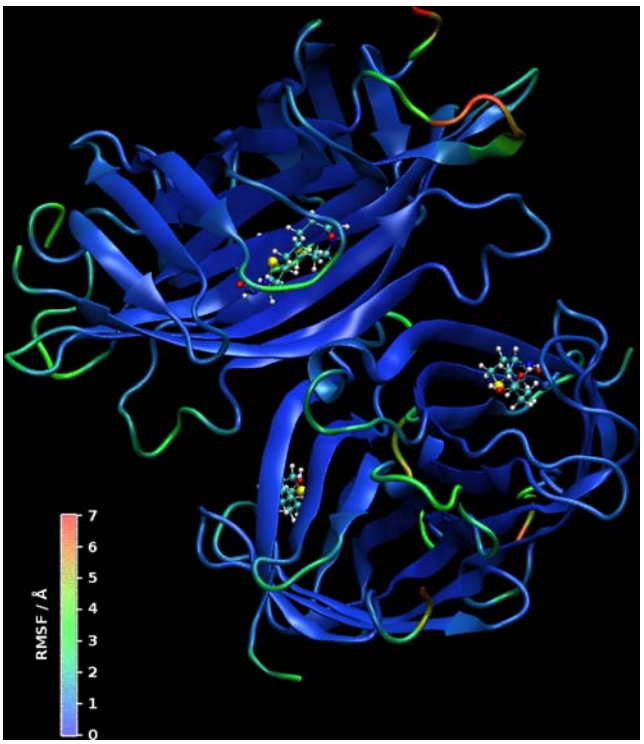

Fig. 6 Root mean square fluctuations (RMSF) of backbone atoms of the studied biotin–WT-SAV complex

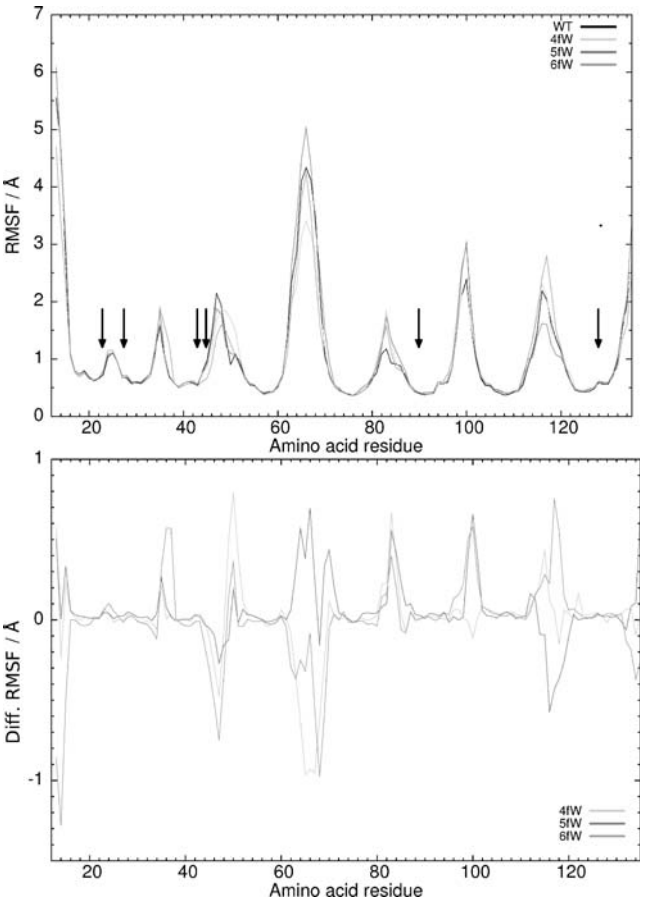

Fig. 7 RMSF of the backbone atoms of individual residues of WT-SAV and its mutants (*upper graph*). The differential RMSF calculated for fluorinated variants (with WT as a reference) is also presented (*lower graph*). Data are averaged for the tetramer

Table 1 Selected interatomic distances corresponding to the biotin–streptavidin (SAV) interaction network (see Fig. 1). Results of the molecular dynamics (MD) simulation. Mean values and standard deviations (in Å) are calculated from the 40 ns simulation for all four binding pockets. *WT* Wild-type SAV; *4fW*, *5fW*, *6fW* 4-, 5- and 6-fluorotryptophan, respectively

Distance	WT	4fW	5fW	6fW
(biotin)O···O(Y43)	2.770±0.190	2.776±0.195	2.761±0.177	2.758±0.160
(biotin)O···O(S27)	2.793±0.260	2.799±0.240	2.833±0.320	2.793±0.265
(biotin)O···N(N23)	3.644±0.542	4.162±0.660	4.041±0.700	4.023±0.684
(biotin)N···O(D128)	4.578±1.296	5.561±1.142	5.043±1.459	5.199±1.341
(biotin)N···O(S45)	3.741±1.010	3.547±0.865	3.277±0.685	3.213±0.449
(biotin)S···O(T90)	3.418±0.266	3.590±0.360	3.542±0.334	3.394±0.249

their description within the non-electronic-structure framework of molecular mechanics is not very reliable.

The dynamics of the hydrogen bonds formed by the carbonyl oxygen atom of the biotin ureido ring with Y43 residue is presented in Fig. 8. It is immediately clear that this O–H···O bond is highly preserved in all the variants throughout the simulation. Events where O···O separation is larger than 3.5 Å are very short and the bond is immediately restored. Interestingly, both 5fW and 6fW variants seem to preserve this contact even better than WT-SAV. This fact is also indicated in Table 1, which groups statistical data for the studied interactions, by shortening the average (biotin)O···O(Y43) distance resulting in smaller standard deviations for 5fW- and 6fW-SAV. The situation is not so clear for the remaining contacts of the biotin ureido oxygen atom; however, the 4fW variant seems to be the most unstable among the studied systems. The flexibility of the O···H–N(Asn23) contact is visibly greater than the two O···H–O contacts, consistent with the generally weaker nature of N–H···O interactions with respect to O–H···O

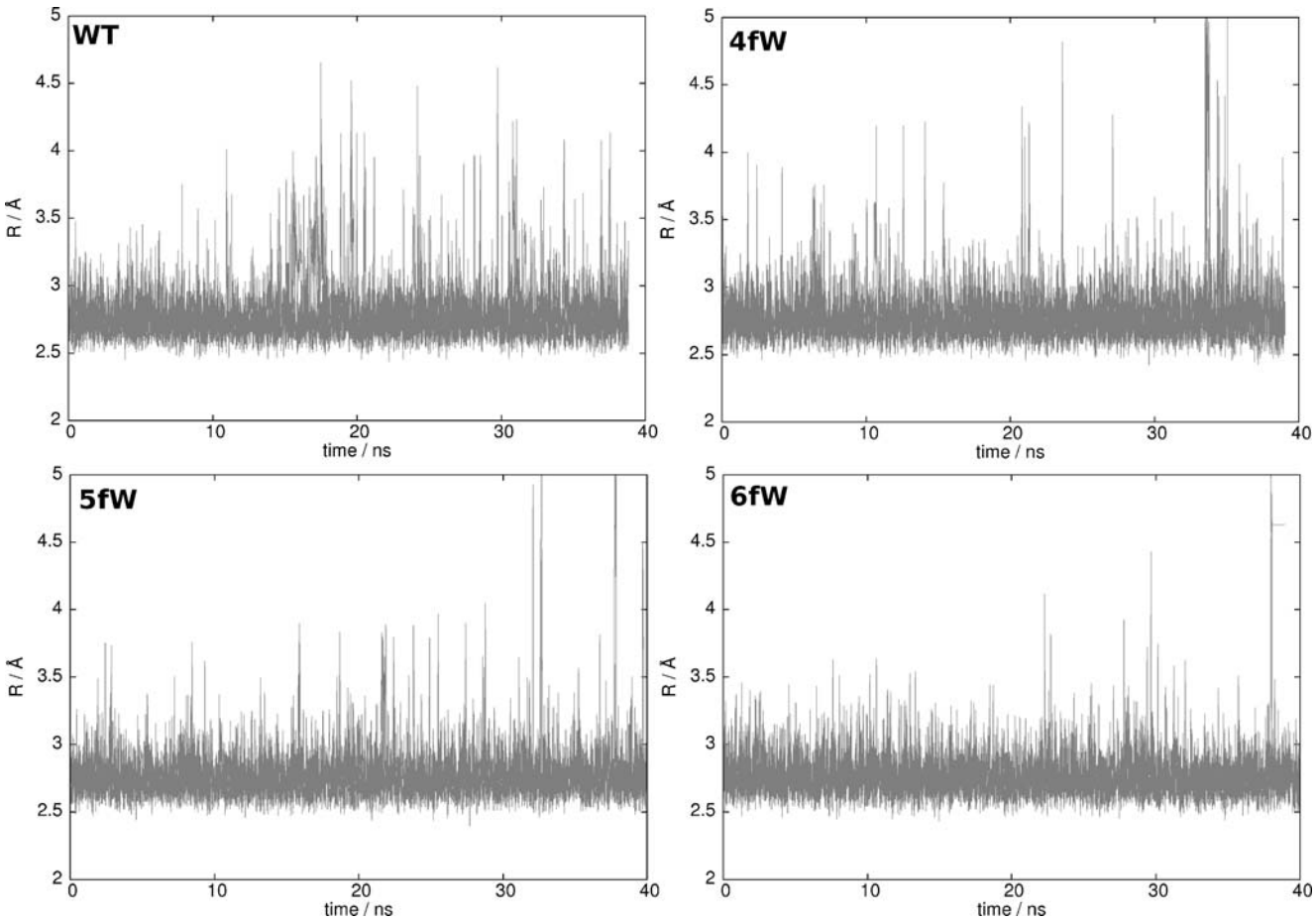

Fig. 8 The O···O interatomic distance (in Å) for the interaction of the biotin residue ureido oxygen atom with the hydroxyl oxygen of the Tyr43 residue as a function of MD simulation time (in ns) for WT-SAV and its fluorotryptophan variants

Table 2 Total interaction energy between biotin and streptavidin together with electrostatic and van der Waals components. Values (in kcal/mol) are the average (± standard deviation) from the MD runs, averaged over all four biotin-binding sites

	WT	4fW	5fW	6fW
Electrostatic energy	−34.48±6.25	−29.95±4.00	−33.31±6.63	−31.54±3.94
Van der Waals energy	−30.39±1.72	−30.04±1.54	−30.22±1.73	−30.82±1.69
Total interaction energy	−64.87±7.22	−59.99±4.67	−63.53±7.67	−62.36±4.77

hydrogen bonds [52]. A similar situation is found for the imine contacts of the ureido ring. One of these, N–H···O (Ser45), is well conserved throughout the simulation, again with 5fW- and 6fW-SAV variants being more stable than the WT complex. The opposite edge of the ureido ring is much more weakly bound, with 5fW and 6fW mutants in the middle between WT (lower limit) and 4fW (largest distance). This fact is significant in the light of a recent elegant study by DeChancie and Houk [53] on the origins of strong, subpicomolar biotin–(strept)avidin binding, revealing the key role of charged residues for inducing and enhancing cooperativity of the hydrogen bond network by polarising the ureido ring of the biotin. The authors identified charged aspartate residues as the primary sources of cooperativity. Our investigations support these findings by revealing the dynamic nature of the hydrogen bonds in the binding pocket. This might also serve as an additional explanation for the observed discrepancy between the experimental and calculated influence of the Asp128 residue on binding enthalpy (see Table 2 in Ref. [53]), since the key stabilising contacts seem not to be available at all times during the simulation. Moreover, in view of our results, which indicate a possible increase in interaction strength for selected contacts in biotin-5fW- and 6fW-SAV (see Table 1), full characterisation of biotin–xfW-SAV systems, in addition to their potential application in extending existing biotin–avidin technology, may allow validation of the model of DeChancie and Houk [53]. Finally, the S(biotin)···HO(Thr90) hydrogen bond exhibits intermediate behaviour between O···O and N···O contacts: it is well conserved throughout the simulation (note the relatively low standard deviations), but its average values are close to the N···O results. The structurising role of the biotin sulphur atom might also be extended beyond the hydrogen-bonded structural stabilisation. The atom is "softer", i.e. more polarisable, than the second-row elements of the ureido moiety, therefore it could also add significantly to the nonpolar binding energy.

Table 2 together with Figs. 9 and 10 describe decomposition of the biotin–streptavidin interaction energy (averaged over four independent binding sites) into electrostatic and van der Waals components. The stability of the non-electrostatic van der Waals energy, which does not change significantly from variant to variant, indicates that the shape of the cavity is indeed well-conserved, and biotin is able to adapt itself to the modified environment. Fluorination does not provide additional van der Waals stabilisation. On the contrary, its influence on the electrostatic component is slightly destabilising. Again, the 4fW-SAV variant has the poorest ability to interact with biotin of the structurally modified proteins studied.

The distribution of electrostatic potential (ESP) around the WT apo-SAV, calculated using the APBS package, is presented in Fig. 11. The entrance to the biotin-binding cavity is clearly visible. While there is a region of positive

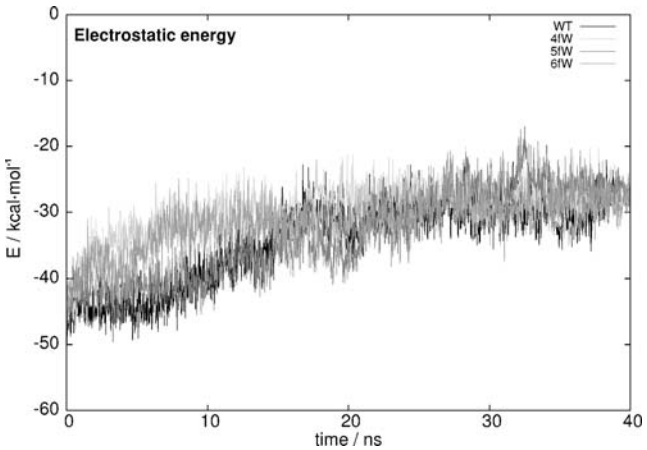

Fig. 9 The electrostatic Coulomb component of the biotin-SAV interaction energy calculated for WT-SAV and its variants as a function of time. Values are averaged over all biotin-binding sites

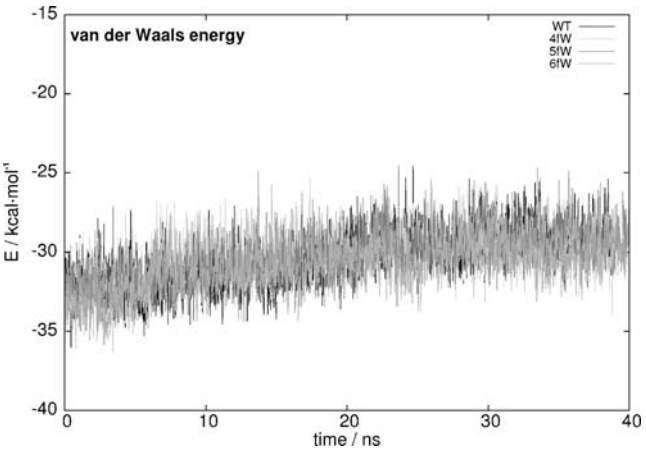

Fig. 10 The van der Waals component of the biotin-SAV energy calculated for WT-SAV and its variants as a function of time. Values are averaged over all biotin-binding sites

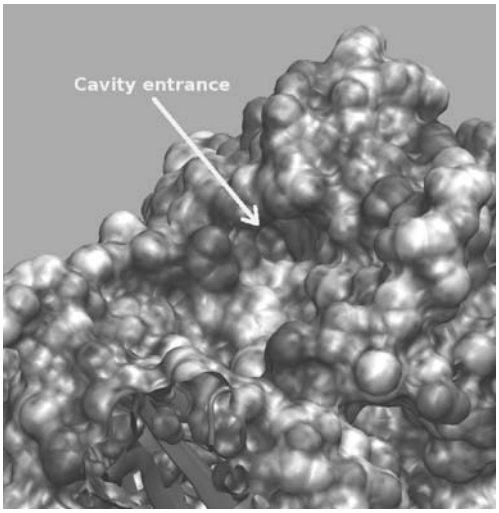

Fig. 11 Electrostatic potential (ESP) distribution in the vicinity of the biotin-binding pocket (biotin not present)

ESP at the entrance to the cavity, negative potential dominates inside the pocket. The cage is accessible to the solvent, and indeed the crystallographic structures of apo-SAV contain either water molecules or salt ions in the binding pocket. Therefore, additional analysis of the solvent-accessible surface (SASA) was carried out (Fig. 12) to provide information on the impact of fluorination on solvent accessibility. All the fluorinated variants tested exhibit increased SASA values throughout the simulation, and the effect is most prominent for the 4fW variant. However, the increase with respect to WT (ca. 200 Å^2) is not large, and is influenced not only by the binding pocket modification, but also by small changes in the overall structure, discussed earlier in terms of the RMSD and RGYR parameters. We conclude that the structural changes to the binding pocket effected by tryptophan fluorination are rather small.

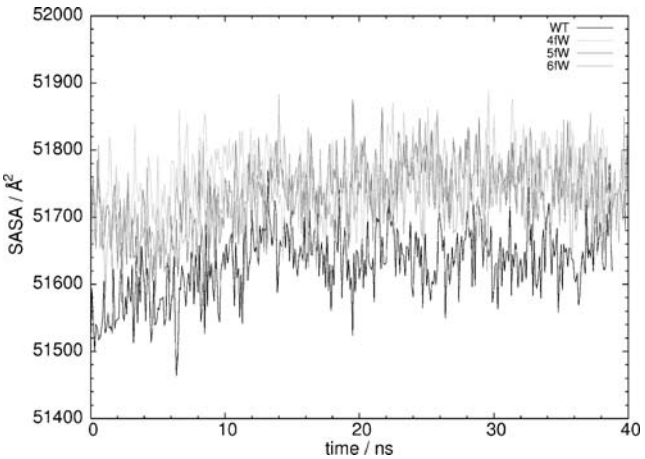

Fig. 12 Solvent accessible surface area (SASA) calculated for WT-SAV and its variants.

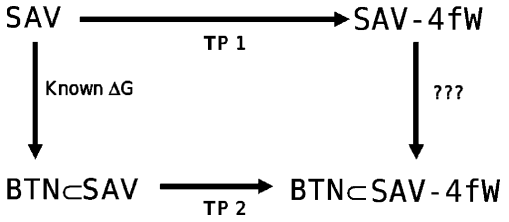

Fig. 13 Thermodynamic perturbation cycle used for the estimation of biotin binding free energy change (ΔΔG) upon tryptophan residue fluorination

The protein stability study indicates that in silico fluorination of tryptophan residues does not destabilise the existing fold. One can assume that the xfW-SAV variants will be stable, with the underlying assumption that the folding process itself will not be affected. The interaction analysis for the ureido ring suggests destabilisation of biotin binding for 4fW, while 5fW- and 6fW-SAV will be less affected or will even bind biotin more strongly. However, this does not take into account other numerous interactions, e.g. differences in hydrophobic effects. An additional thermodynamic perturbation study was therefore performed according to the thermodynamic cycle presented in Fig. 13. Again here, we assume that the biotin-binding/biotin-release to and from SAV, as studied computationally in detail by Izrailev et al. [9], will be affected by tryptophan fluorination *only* when the biotin is firmly docked in the pocket, while the 3–4 loop opening phase and biotin entry event will not be perturbed. The free energy difference of two runs—TP1 (without biotin) and TP2 (with biotin)—then corresponds directly to the ΔΔG (change in the free energy) of biotin binding.

The forward and backward alchemical transformations (Trp to/from fluorinated Trp) converge within 0.5 kcal/mol (1 kcal/mol for 6fW-SAV), despite rather large hysteresis (visible in Fig. S2—a representative example, biotin-containing WT-SAV to 4fW-SAV transformation). Calculated ΔΔG values (Table 3) show that tryptophan fluorination has only a minimal destabilising effect on the free energy of biotin binding. The reported values are calculated for the tetrameric protein containing four biotin molecules, therefore the ΔG change per one biotin molecule does not exceed +1.7 kcal/mol, provided that fluorination does not change the folding of the protein or the biotin entrance pathway. All the contributions are destabilising, but they closely follow earlier findings

Table 3 Thermodynamic perturbation results for biotin binding. Free energies in kcal/mol for the tetrameric system

Mutation	4fW-SAV	5fW-SAV	6fW-SAV
TP2 (with four biotin ligands)	−117.0	−129.0	−292.5
TP1 (no biotin ligands)	−123.5	−132.0	−297.5
Total ΔΔG	+6.5	+3.0	+5.0

based on MD trajectories. The 5fW-SAV is the least affected species, with the 6fW and 4fW variants gradually losing biotin binding ability. Remarkably, the 6fW substitution is predicted to have a much more pronounced energy effect on the protein itself than the other two mutations. The loss of binding force is comparable to that reported by Stayton and coworkers [19] for conservative mutations of tryptophan to phenylalanine. We therefore expect that, with careful design of biotin-like ligands, one could achieve binding affinities similar to, if not better than, those of the prototypic SAV–biotin system.

Conclusions

The computational study presented here describes the effects of tryptophan residue fluorination on both structure and biotin-binding abilities of streptavidin. Overall, the structural and energetic effects are predicted to be small, which is promising in terms of both protein stability and reversible binding applications. Consistently, the 4fW-substituted variant displays the most reduced biotin-binding ability, and the most disrupted hydrogen bonding pattern between the ligand and the host. The 5fW-SAV variant is potentially least affected by fluorination, and there are indications that, for this species, the hydrogen-bonding properties of the binding pocket are enhanced. This protein is the tentative target for further studies with other biotinylated and analogous ligands, which may not only extend existing biotin–avidin technology, but may also facilitate validation of theoretical models of the origin of subpicomolar binding.

Acknowledgements J.J.P. would like to express his gratitude to the group of Prof. Qiang Cui, Department of Chemistry, University of Wisconsin, especially Qiang Cui, Peter Koenig and Xavier Prat-Resina, for their friendly support. The CPU time grants from Wrocław Centre for Networking and Supercomputing (WCSS) and Poznan Supercomputing and Networking Centre (PCSS) are gratefully acknowledged. This work is supported by the European Union as the FP6 Marie Curie Research Training Network IBAAC Project, MRTN-CT-2003-505020, and by the Slovenian Ministry of Higher Education, Science and Technology research grant P1-017.

References

1. Savage D (1993) Avidin-biotin chemistry: a handbook. Pierce, Rockford, IL
2. Wilchek M, Bayer EA (1990) Methods Enzymol 184:5–13. doi:10.1016/0076-6879(90)84256-G
3. Green NM (1975) Adv Protein Chem 29:85–133. doi:10.1016/S0065-3233(08)60411-8
4. Weber PC, Wendoloski JJ, Pantoliano MW, Salemme FR (1992) J Am Chem Soc 114:3197–3200. doi:10.1021/ja00035a004
5. Rao J, Lahiri J, Isaacs L, Weis RM, Whitesides GM (1998) Science 280:708–711. doi:10.1126/science.280.5364.708
6. Livnah O, Bayer EA, Wilchek M, Sussman JL (1993) Proc Natl Acad Sci USA 90:5076–5080. doi:10.1073/pnas.90.11.5076
7. Klumb LA, Chu V, Stayton PS (1998) Biochemistry 37:7657–7663. doi:10.1021/bi9803123
8. Weber PC, Ohlendorf DH, Wendoloski JJ, Salemme FR (1989) Science 243:85–88. doi:10.1126/science.2911722
9. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K (1997) Biophys J 72:1568–1581
10. González M, Bagatolli LA, Echabe I, Arrondo JLR, Argaraña CE, Cantor CR, Fidelio GD (1997) J Biol Chem 272:11288–11294. doi:10.1074/jbc.272.1.497
11. Houk KN, Leach AG, Kim SP, Zhang X (2003) Angew Chem Int Ed 42:4872–4897. doi:10.1002/anie.200200565
12. Klein G, Humbert N, Gradinaru J, Ivanova A, Gilardoni F, Rusbandi UE, Ward TR (2005) Angew Chem Int Ed 44:7764–7767. doi:10.1002/anie.200502000
13. Loosli A, Rusbandi UE, Gradinaru J, Bernauer K, Schlaepfer CW, Meyer M, Mazurek S, Novič M, Ward TR (2006) Inorg Chem 45:660–668. doi:10.1021/ic051405t
14. Laitinen OH, Hytönen VP, Nordlund HR, Kulomaa MS (2006) Cell Mol Life Sci 63:2992–3017. doi:10.1007/s00018-006-6288-z
15. Sano T, Vajda S, Cantor CR (1998) J Chromatogr B Biomed Sci Appl 715:85–91. doi:10.1016/S0378-4347(98)00316-8
16. Malmstadt N, Hyre D, Ding Z, Hoffman AS, Stayton PS (2003) Bioconjug Chem 14:575–580. doi:10.1021/bc0200551
17. Kuhn B, Kollman PA (2000) J Am Chem Soc 122:3909–3916. doi:10.1021/ja994180s
18. Dixon RW, Radmer RJ, Kuhn B, Kollman PA, Yang J, Raposo C, Wilcox CS, Klumb LA, Stayton PS, Behnke C, Le Trong I, Stenkamp R (2002) J Org Chem 67:1827–1837. doi:10.1021/jo991846s
19. Chilkoti A, Tan PH, Stayton PS (1995) Proc Natl Acad Sci USA 92:1754–1758. doi:10.1073/pnas.92.5.1754
20. Chilkoti A, Stayton PS (1995) J Am Chem Soc 117:10622–10628. doi:10.1021/ja00148a003
21. Miyamoto S, Kollman PA (1993) Proc Natl Acad Sci USA 90:8402–8406. doi:10.1073/pnas.90.18.8402
22. Sixl F, King RW, Bracken M, Feeney J (1990) Biochem J 266:545–552
23. Li E, Quian SJ, Nader L, Yang NC, d'Avignon A, Sacchettini JC, Gordon JI (1989) J Biol Chem 264:17041–17048
24. Li E, Locke B, Yang NC, Ong DE, Gordon JI (1987) J Biol Chem 262:13773–13779
25. Clarkson J, Smith DA, Masca S, Batchelder DN (2001) Biopolym Biospectr 62:307–314
26. Wang X, Mercier P, Letourneau P-J, Sykes BD (2005) Protein Sci 14:2447–2460. doi:10.1110/ps.051595805
27. Fleischman SH, Brooks CL III (1987) J Chem Phys 87:3029–3037. doi:10.1063/1.453039
28. Katz BA (1997) J Mol Biol 274:776–800. doi:10.1006/jmbi.1997.1444
29. Becke AD (1993) J Chem Phys 98:5648–5652. doi:10.1063/1.464913
30. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785–789. doi:10.1103/PhysRevB.37.785
31. Besler BH, Merz KM Jr, Kollman PA (1990) J Comput Chem 11:431–439. doi:10.1002/jcc.540110404
32. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr. JA, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J,

Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski V.G, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian03, Rev. C.02, Gaussian, Wallingford, CT

33. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) J Phys Chem B 102:3586–3616. doi:10.1021/jp973084f
34. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935. doi:10.1063/1.445869
35. Darden T, York D, Pedersen LG (1993) J Chem Phys 98:10089–10092. doi:10.1063/1.464397
36. Essman U, Perela L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577–8592. doi:10.1063/1.470117
37. Sagui C, Darden T (1999) Annu Rev Biophys Biomol Struct 28:155–179. doi:10.1146/annurev.biophys.28.1.155
38. Adelman SA, Doll JD (1976) J Chem Phys 64:2375–2388. doi:10.1063/1.432526
39. Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) J Chem Phys 103:4613–4621. doi:10.1063/1.470648
40. Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341. doi:10.1016/0021-9991(77)90098-5
41. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) J Comput Chem 26:1781–1802. doi:10.1002/jcc.20289
42. Humphrey W, Dalke A, Schulten K (1996) J Mol Graph 14:33–38. doi:10.1016/0263-7855(96)00018-5
43. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) Nucleic Acids Res 32:W665–W667. doi:10.1093/nar/gkh381
44. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Proc Natl Acad Sci USA 98:10037–10041. doi:10.1073/pnas.181342398
45. Le Trong I, Freitag S, Klumb LA, Chu V, Stayton PS, Stenkamp RE (2003) Acta Crystallogr D Biol Crystallogr 59:1567–1573. doi:10.1107/S0907444903014562
46. Hyre DE, Le Trong I, Merritt EA, Eccleston JF, Green NM, Stenkamp RE, Stayton PS (2006) Protein Sci 15:459–467. doi:10.1110/ps.051970306
47. Marttila AT, Hytönen VP, Laitinen OH, Bayer EA, Wilchek M, Kulomaa MS (2003) Biochem J 369:249–254. doi:10.1042/BJ20020886
48. Babu MM, Singh SK, Balaram P (2002) J Mol Biol 322:871–880. doi:10.1016/S0022-2836(02)00715-5
49. Aravinda S, Shamala N, Bandyopadhyay A, Balaram P (2003) J Am Chem Soc 125:15065–15075. doi:10.1021/ja0372762
50. Klaholz BP, Moras D (2002) Structure 10:1197–1204. doi:10.1016/S0969-2126(02)00828-6
51. Sarkhel S, Desiraju GR (2004) Proteins. Struct Funct Bioinf 54:247–259. doi:10.1002/prot.10567
52. Rudner MS, Jeremic S, Petterson KA, Kent DR, Brown KA, Drake MD, Goddard WA, Roberts JD (2005) J Phys Chem A 109:9076–9082. doi:10.1021/jp052925c
53. DeChancie J, Houk KN (2007) J Am Chem Soc 129:5419–5429. doi:10.1021/ja066950n