

## MODELING OF THREE DIMENSIONAL STRUCTURE OF HUMAN ALPHA-FETOPROTEIN COMPLEXED WITH DIETHYLSTILBESTROL: DOCKING AND MOLECULAR DYNAMICS SIMULATION STUDY

ALEXANDER A. TEREENTIEV<sup>\*,‡</sup>, NURBUBU T. MOLDOGAZIEVA<sup>\*,§</sup>,  
OLGA V. LEVTSOVA<sup>†</sup>, DMITRY M. MAXIMENKO<sup>\*</sup>, DENIS A. BORODZENKO<sup>\*</sup>  
and KONSTANTIN V. SHAITAN<sup>†,¶</sup>

<sup>\*</sup>*Department of Biochemistry, Russian State Medical University  
1, Ostrovyanova Street, Moscow 117997, Russia*

<sup>†</sup>*Department of Bioengineering  
M.V. Lomonosov Moscow State University  
1, Vorobyevy Gory, Moscow 119991, Russia*

<sup>‡</sup>*aaterent@inbox.ru*

<sup>§</sup>*nmoldogazieva@mail.ru*

<sup>¶</sup>*shaitan@moldyn.org*

Received 15 October 2011

Revised 28 January 2012

Accepted 28 January 2012

Published 12 April 2012

It has been long experimentally demonstrated that human alpha-fetoprotein (HAFP) has an ability to bind immobilized estrogens with the most efficiency for synthetic estrogen analog — diethylstilbestrol (DES). However, the question remains why the human AFP (HAFP), unlike rodent AFP, cannot bind free estrogens. Moreover, despite the fact that AFP was first discovered more than 50 years ago and is presently recognized as a “golden standard” among onco-biomarkers, its three-dimensional (3D) structure has not been experimentally solved yet. In this work using MODELLER program, we generated 3D model of HAFP on the basis of homology with human serum albumin (HSA) and Vitamin D-binding protein (VTDB) with subsequent molecular docking of DES to the model structure and molecular dynamics (MD) simulation study of the complex obtained. The model constructed has U-shaped structure in which a cavity may be distinguished. In this cavity the putative estrogen-binding site is localized. Validation by RMSD calculation and with the use of PROCHECK program showed good quality of the model and stability of extended region of four alpha-helical structures that contains putative hormone-binding residues. Data extracted from MD simulation trajectory allow proposing two types of interactions between amino acid residues of HAFP and DES molecule: (1) hydrogen bonding with involvement of residues S445, R452, and E551; (2) hydrophobic interactions with participation of L138, M448, and M548 residues. A suggestion is made that immobilization of the hormone using a long spacer provides delivery of the estrogen molecule to the binding site and, thereby, facilitates interaction between HAFP and the hormone.

**Keywords:** Homology-based modeling; alpha-fetoprotein; serum albumin; Vitamin D-binding protein; estrogens; complex with diethylstilbestrol; molecular docking, molecular dynamics.

## 1. Introduction

Alpha-fetoprotein (AFP) is the major mammalian tumor-associated fetal protein that is normally present in the blood of adults only in small quantities (up to 10–20 ng/ml).<sup>1,2</sup> AFP has been recognized as a growth regulator during embryonic development and tumor progression.<sup>3</sup> It has been shown that AFP in dose-dependent manner stimulates growth of estrogen-resistant tumors and inhibits growth of estrogen-sensitive ones.

Chemically, AFP is a glycoprotein with molecular weight (MW) of about 68–73 kDa and carbohydrate content of 3%–5%. Primary structure of full-length human AFP (HAFP) molecule is represented by 609 amino acid (aa) residues from which a signal peptide of 18 aa residues is removed during processing to yield the mature protein of 591 aa residues.<sup>4,5</sup>

AFP belongs to the family of proteins — products of albuminoid genes, which are located in tandem arrangement in chromosome 4 (region 4q11–q13).<sup>6</sup> This family includes (Table 1), besides AFP itself, also serum albumin (SA), Vitamin D-binding protein (VTDB) and alpha-albumin (afamin). Proteins of this family are evolutionary closely related and originate from the common ancestor.<sup>7</sup> They demonstrate considerable similarity of their primary structures with characteristic distribution of disulphide bonds.

AFP and SA show the highest degree of identity (up to 40%) and similarity (taking into account conservative substitutions of aa residues) up to 60% amongst the family members. Proteins of this family also demonstrate similar  $\alpha$ -helical secondary structure (for example, up to 65%–67%  $\alpha$ -helices in AFP and about 50% in SA). It has been shown using electron microscopy, image processing and circular dichroism that AFP and SA have U-shaped spatial organization and consist of three homologous domains (I, II and III).<sup>8</sup> Each domain is composed of two subdomains that are designated as IA, IB, IIA, IIB, IIIA and IIIB and possess common structural motifs. It has been also shown that the principal regions of ligand-binding in human serum albumin (HSA) are located in hydrophobic cavities in subdomains IIA and IIIA, which exhibit similar chemistry.<sup>9–11</sup>

As for 3D structure of proteins of this family, numerous crystal structures of human SA and VTDB free of ligands or complexed with variety of low molecular weight hydrophobic ligands have been obtained.<sup>12–16</sup> To date (December 2011), 86 crystal structures of HSA itself and in complexes with fatty acids, drugs, bilirubin, thyroxine and heme at resolutions between 1.90–3.20 Å are present in the Protein Data Bank (PDB).<sup>17</sup> Also six crystal structures of VTDB uncomplexed or in complexes with vitamin D and its metabolites and also with skeletal actin are determined and exist in PDB.

However, the precise 3D structure of HAFP has not been experimentally obtained yet. This is because of difficulties in crystallization of this protein that impede study of 3D structure of HAFP by experimental procedures such as X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy.

Table 1. Properties of proteins — products of albuminoid gene family.

Name of the protein	MW, kDa	Full length, (aa numbers in mature protein)	Signal peptide length, aa amount	Amount of S-S bonds	Carbohydrate content, %	Domain composition	Functions
Alpha-fetoprotein	68–73	609 (aa res. 19–609)	18	15	3–5	I (19–210), II(211–402), III(403–601) in human	Specific binding of hydrophobic ligands. Regulation of cell proliferation and immune response
Serum albumin	66–68	609 (aa res. 25–609)	18 + 6 (propep.)	17	0.5	I (19–210), II(211–403), III(404–601) in human	Binding and transportation of variety of hydrophobic ligands
Vitamin D-binding protein	52–58	474 (aa res. 17–474)	16	14	5.0	I (17–208), II(209–394), III(395–474) in human	Transportation of Vitamin D metabolites, fatty acids and skeletal actin
Afamin	82–87	599 (aa res. 22–599)	21	17	21	I (22–210) II(211–403) III(404–599)	Binding of Vitamin E, prevention of nervous cells from apoptosis

Identity degree (40%) between primary structures of AFP and SA is quite enough to make it possible comparative modeling of three-dimensional structure of AFP on the basis of homology with SA.<sup>18</sup> Homology-based modeling is a computer-based approach that employs the fact that spatial structures of homologous proteins are evolutionary more conserved than their amino acid sequences provided by identity degree higher than 30%.<sup>19</sup> This method allows constructing an atomic resolution structural model of a “target” protein from its amino acid sequence that is aligned with that of homologous “template” protein of known, experimentally obtained, 3D structure. It can provide models useful for generating hypotheses about functions of proteins and the role of conserved aa residues in binding of low-molecular-weight ligands.

It has been long demonstrated that AFP from different sources binds and, seemingly, transports various hydrophobic ligands including estrogens, fatty acids, bilirubin, some drugs and dyes.<sup>20,21</sup> Studying of estrogen-binding ability of AFP is of crucial importance since this ability provides regulation of free active forms of hormones in conditions *in vivo*. In this way AFP is supposed to protect embryonic tissues from circulating maternal estrogens. However, binding of natural free estrogens has been evidently demonstrated only for mouse and rat alpha-fetoproteins (MAFP and RAFP, respectively). For HAFP, an ability to bind immobilized estrogens with the highest efficiency for synthetic analog of estrogens — diethylstilbestrol (DES) — has been shown.<sup>22</sup> In the latter work, HAFP was isolated from embryonic tissues by extraction with butanol that leads to delipidation of HAFP molecule to make its estrogen-binding sites free from hormones and fatty acids. Also, a suggestion has been made that immobilization is required for proper orientation of the hormone molecule that is necessary for its binding to HAFP molecule.

In order to identify amino acid residues required for estrogen binding, 17 mutant human AFPs were produced by site-directed mutagenesis.<sup>23</sup> In these mutant HAFPs some amino acid residues were substituted with those in RAFP at corresponding positions. Substitutions of amino acid residues in the domain III, namely, A455V, A456S, T457I, L469R, L470S, A476L and I480Y resulted in obtaining of HAFP molecules that showed significant ability to bind free estrogens. The conclusion was made that aa residues V455, S456, I457, R469, S470, L476, Y480 should be at the corresponding positions of HAFP to convert it into the form that able to bind estrogens.

Using native and recombinant chimeric rat-human proteins it has been identified that estrogen-binding region of rodent AFP encompasses amino acid residues from 423 to 506 located in the domain III.<sup>24</sup> In this region the peptide segment composed of 15 aa residues with sequence ELIDLTGKMVSIAS (aa 448–462 in rat AFP and aa 442–456 in mouse AFP) has been proposed to be a primary high-affinity (Kd about  $10^{-8}$  M) estrogen-binding site of rodent AFP (Table 2). In human AFP, the peptide motif composed of aa residues 446–460 with sequence ELMAITRKMAA-TAAT, has been suggested to be the primary high-affinity estrogen-binding site on the basis of amino acid matching with rodent AFP.<sup>25</sup> Besides, synthetic peptide

Table 2. Alignment of amino acid residues of primary high affinity ( $K_d = 10^{-8} \text{ M}^{-1}$ ) estrogen-binding sites of human, mouse and rat AFPs.

Biological species	Amino acid sequence	Numbers of aa residues of the sequence	Identity degree	Similarity degree	Accession number to UniProtKB/Swiss-Prot Data base
human	ELMAITRKMAATAAT ** : * * * : * *	446–460	—	—	P02771
rat	ELIDLTGKMVSIAS ** : * * * : * *	448–462	7/15 (46.7%)	10/15 (66.7%)	P02773
mouse	ELIDLTGKMVSIAS ** : * * * : * *	442–456	7/15 (46.7%)	10/15 (66.7%)	P02772

Note: (\*) indicates identity, (:) — conservative substitutions.

derived from HAFP, namely AADIIIGHLCIRHEM (aa residues 476–490), has been experimentally demonstrated to bind  $17\beta$ -estradiol (E2) with moderate affinity ( $K_d$  about  $10^{-5} \text{ M}$ ).<sup>26</sup> The latter sequence was proposed to be a secondary, low-affinity estrogen-binding site. So, it may be concluded that estrogen binding by HAFP should take place with involvement of amino acid residues which are included in the region 446–490 in the full-length molecule.

Nevertheless, the precise molecular mechanisms of estrogen binding to HAFP and the role of definite aa residues in this process have not been elucidated yet. In this work we generated the 3D structure of HAFP on the basis of homology with HSA and VTDB followed by docking of synthetic estrogen analog diethylstilbestrol. Subsequent molecular dynamics (MD) simulation study was performed to optimize the model and to elucidate amino acid residues that, putatively, participate in the protein–hormone interaction.

## 2. Materials and Methods

### 2.1. Template selection and sequence alignment

Homology-based modeling is an effective method for predicting 3D structure of a protein, provided by experimentally obtained 3D structure of homologous protein. The latter can serve as a template molecule for model generation. Correct template selection in homology-based modeling is one of the most important factors that provide high quality of the model constructed.<sup>27</sup>

Here candidate templates for human AFP were searched from PDB database. To identify the best template the following criteria were used: close relationship to AFP, similar functions, relatively high degree of sequence identity, experimentally observed secondary structure similarity, and also high resolution of experimentally obtained 3D structure of the template molecule. Using these approaches two 3D structures as templates were chosen from PDB data base — those of human serum albumin (accession number in PDB 1E78) and VTDB (accession number in PDB 1KW2). Crystal structures for both proteins were determined by X-ray diffraction method at high resolutions: 2.6 Å for HSA and 2.15 Å for VTDB.

Amino acid sequences of the template molecules were aligned to the sequence of human AFP. The sequences of HSA (P02768), human VTDB (P02774) and HAFP (P02771) were retrieved in FASTA format from NCBI (National Center for Biotechnology Information) UniProtKB/Swiss-Prot database (<http://www.uniprot.org/>).<sup>28</sup> Sequence alignment was performed using the program for global multiple alignment ClustalW, version 2.0.<sup>29</sup>

## 2.2. Model generation, optimization and validation

The 3D model of human alpha-fetoprotein was generated using the MODELLER 9v6 program package for comparative modeling and protein structure prediction.<sup>30</sup> This is the most commonly used software that utilizes spatial-restraint method to construct 3D structure by calculating positions of all heavy atoms in a target protein from data experimentally obtained for template protein molecule. Restraints are applied to the main protein internal coordinates — protein backbone distances and dihedral angles. They serve as the basis for global optimization procedure by conjugate energy minimization.<sup>31,32</sup> To perform optimization and relaxation of the model constructed and to assess its stability MD simulation at the trajectory length of 5 ns was carried out.

Assessment of the model accuracy was performed using the following two methods. The first one was comparing two protein structures — those of initial model constructed and of the model structure during 5 ns MD simulation. This method uses the root-mean-square deviation (RMSD) metric to measure the mean distances between the corresponding atoms in the two structures after they have been superimposed.<sup>33,34</sup> In this work, calculation of RMSD values were performed for four extended alpha-helices located in the cavity revealed in the model structure. This region contains hypothetical estrogen-binding site and was of our special interest to estimate possibility for docking of the estrogen molecule to HAFP.

The second method used for validation of our 3D model was PROCHECK, version 3.5, program that verified the conformational and stereochemical parameters of amino acid residues of a protein. These parameters are: combinations of dihedral angles  $\varphi$  and  $\psi$  values, peptide bond planarity, main chain and side chain parameters, main chain covalent bond lengths and angles, residue-by-residue stereochemical properties such as absolute deviation of torsion angles and C $\alpha$ -chirality.<sup>35,36</sup> Also, it allowed receiving information about secondary structure elements and estimated solvent accessible areas. Judgment about how “normal” or “unusual” is a given stereochemical property was done by measuring *G*-factor. A low (or negative) *G*-factor indicates that the property corresponds to a low-probability conformation. So, PROCHECK program allows judging about a model quality by revealing poor and unusual geometry along a protein sequence.

## 2.3. Construction of DES molecule

Synthetically obtained analog of estrogens DES was used as a ligand for docking. Construction of the ligand molecule was performed with the use of HyperChem

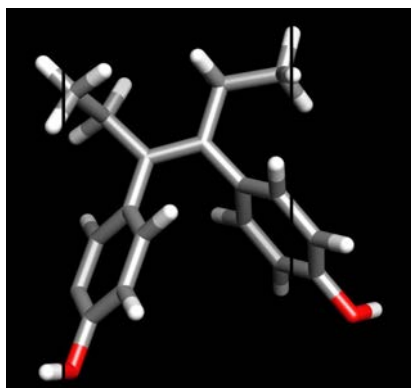


Fig. 1. Structure of diethylstilbestrol after geometry optimization. The structure was constructed with the use of HyperChem (version 8.0) molecular modeling software.

(version 8.0) molecular modeling software.<sup>37</sup> Modeling of the hormone molecule (Fig. 1) was carried out in the full-atomic approximation in Amber potential force field.<sup>38,39</sup> The force field parameters were extended with experimental data obtained by quantum chemistry calculations performed with the use of GAMESS software package.<sup>40</sup> Effective charges on atoms were calculated using Mulliken method.<sup>41</sup> The hormone molecule constructed was placed in the periodic box of  $20 \times 20 \times 20 \text{ \AA}$  in size with explicit water environment. Geometry optimization of the ligand structure was performed by energy minimization with the use of restricted Hartree–Fock method and the standard basis set 6-31GF.<sup>42</sup>

#### 2.4. Docking of DES to human AFP model constructed

Using of 3D structure constructed by homology-based modeling approach for molecular docking is possible because functional regions of a protein, especially its binding sites, tend to be more highly conserved and thus more accurately modeled.<sup>43</sup>

Program package AutoDock (version 4.0)<sup>44</sup> was used to perform automated docking of the estrogen molecule to binding site of HAFP model obtained. In this study, semi-flexible docking was performed where the protein and ligand structures were kept as rigid, but amino acid residues that are hypothesized to participate in estrogen-binding after preliminary docking were kept flexible. These were the following five aa residues: Q428, M448, R452, V543 and T547.

The protocol for this semi-flexible protein-estrogen docking consisted of 10 independent runs per ligand, using an initial population of 100 randomly placed individuals, with  $2.5 \times 10^6$  energy evaluations, a maximum number of 27,000 iterations, a mutation rate of 0.02, and a crossover rate of 0.80. Step sizes were  $0.2 \text{ \AA}$  for translations,  $5.0^\circ$  for quaternions and  $5.0^\circ$  for torsions. For docking simulation, cluster tolerance was set at  $0.5 \text{ \AA}$  and external grid energy was kept at 1000 kcal, with maximum initial energy 0.0 kcal. The probability to perform a local search on an individual

in the population was 0.06, using a maximum of 300 iterations per local search. The best conformers were selected on the basis of Lamarckian Genetic Algorithm (LGA). After docking, the best 10 solutions were clustered into groups with RMSD lower than 1.0 Å. The clusters were ranked by the lowest energy representative of each cluster.

### 2.5. *Molecular dynamics simulation protocol*

The complexes obtained were subjected to molecular dynamics simulation at trajectory length of 5 ns. Two MD simulations were carried out: one for optimization of HAFP model generated, the second for the docked HAFP-DES complex structure. The latter was performed to investigate conformational mobility of the protein backbone and to take information about interaction between amino acid residues of the binding site and the hormone molecule.

MD simulations both in case of relaxation of 3D model of HAFP itself and in case of the HAFP-DES complex were performed using the same parameter sets (protocols). GROMACS software package (version 3.3.1) was utilized.<sup>45</sup> MD simulation was performed in periodic box with sizes  $70 \times 100 \times 100$  Å and with the use of explicit water model. Cutoff distances for electrostatic and van der Waals interactions were given to be equal to 2 nm. LINKS algorithm for fixation of covalent bonds was used. Stochastic dynamics as an algorithm for integration was used to maintain constant temperature in OPLS-AA force field. Time constant for temperature coupling was 0.2 ps. Integration step length was 2 fs. Sodium ( $\text{Na}^+$ ) and chlorine ( $\text{Cl}^-$ ) ions were added to the system to neutralize charges.

## 3. Results

### 3.1. *Overall architecture of the structural model*

Sequence alignment showed 40% identity and 60% overall similarity between HAFP and HSA [Fig. 2(a)]. It is known that a model tends to be reliable, with only minor errors in side chain packing and rotameric state, if identity degree between the template and target proteins is above 50%. Great structural differences arise if identity degree decreases to 20%.<sup>18</sup> So, identity degree between our template and target sequences (40%) is quite enough to perform homology-based modeling with subsequent docking of the hormone molecule, especially if take into account that identity degree is highest (48%) in the domain III, where the putative estrogen-binding site of HAFP is located. Usage of VTDB as an additional template increases reliability of the model constructed in spite of only 21% identity between HAFP and VTDB [Fig. 2(b)]. Low degree of identity between the two latter proteins is due to the fact that VTDB is truncated at its C-terminus and contains only 80 aa residues in its domain III as a result of molecular evolution.<sup>46</sup> Full-length molecules of HAFP and HSA contain 609 aa residues whereas of VTDB — only 474 aa residues are present.

Sequence alignment between HAFP and HSA [Fig. 2(a)] with indication of secondary structure elements shows that there is low amount of flexible loop regions in



Modeling of 3D Structure of Human Alpha-Fetoprotein

Score = 496 bits (1277), Expect = 1e-144, Method: Compositional matrix adjust.  
 Identities = 244/610 (40%), Positives = 363/610 (60%), Gaps = 2/610 (0%)

```

Query 1 MKWVESIFLIFLLNFTESTRTLHRN|EYGIASILDSYQCTAEISLADLATIFFAQFVQEATY 60
MKWV I L+FL + SR + R |+ + + ++ E + L I FAQ++Q+ +
Sbjct 1 MKWVTFISLLFLFSSAYSRSRVFRR|DAHKSEVAHRFKDLGEEFNKALVLIAPQYLQCCPF 60
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 61 KEVSKMVKDALTAIEKPTGDEQSSGLENQLPAFLEELCHEKEILEKYGH-SDCCSQSEE 119
++ K+V + + DE + C ++ F ++LC + E YG +DCC++ E
Sbjct 61 EDHVKLVNEVTEFAKTCVADESAENCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEP 120
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 120 GRHNCFLAHKKPTPASIPLFQVPEPVTSCAEYEDRETFMNKFIYEIARRHPFLYAPTIL 179
R+ CFL HK P ++P PE C A+ ++ ETF+ K++YEIARRHP+ YAP +L
Sbjct 121 ERNECFLQHKDDNP-NLPRLVPEVDVMCTAFHDNEETFLLKLYEYIARRHPYFYAPELL 179
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 180 LWAARYDKIIPSCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAIT 239
+A RY CC+A + C K + E + SS + CA ++ FG R F+A
Sbjct 180 FFAKRYKAAFTGCCQAADKAACLLPKLDELDRDEGKASSAKQRLKCASLQKQGERAFKAWA 239
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 240 VTKLSQKFTKVNFTIQLKLVLDVAHVHEHCCRGDVLDCLQDGEKIMSUYICSQQDTLNSKI 299
V +LSQ+F K F E+ KLV D+ VH CC GD+L+C D + YIC QD++S+K+
Sbjct 240 VARLSQRFPAEFAEVSXLVTDLTQVHTECCGDLLCADDRADLAKYICENQDSISSKL 299
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 300 TECCKLTLTERGQCCIHAENDEKPEGLSPNLNRFLGDRDFNQFSSGKNIIFLASFVHEYS 359
ECC+ LE+ CI ENDE P L F+ +D + + K++FL F++EY+
Sbjct 300 KECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYA 359
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 360 RRHPQLAVSVILRVAKGYQELLEKCFQTENPLECQDKGEEELQKYIQESQALAKRSCGLF 419
RRHP +V ++LR+AK Y+ LEKC +P EC K +E + ++E Q L K++C LF
Sbjct 360 RRHPDYSVVLRLRLAKTYETLEKCCAAADPHECYAKVFDEFKPLVEEPQNLKQNCSELF 419
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 420 QKLGEYYLQNAFLVAYTKKAPQLTSSSELMAITRKMAATAATCCQLSEDKLLACGEGAADI 479
++LGEY QNA LV YTKK PQ+++ L+ ++R + + CC+ E K + C E +
Sbjct 420 EQLGEYKFNALLVRYTKVPQVSTPTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSV 479
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 480 IIGHLCIRHEMTPVNPVGVQCCTSSYANRRPCFSSLVVDETYVPPAFSDDKFIHFKDLCQ 539
++ LC+ HE TPV+ V +CCT S NRRPCFS+L VDETYVP F+ + F FH D+C
Sbjct 480 VLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETTFHADICT 539
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 540 AQGVALQTMKQEFILINLVKQKQPITEEQLEAVIADFSGLLEKCCQGQEQEVCFAEEGQKL 599
Q KQ L+ LVK KP+ T+EQL+AV+ DF+ +EKCC+ ++E CFAEEG+KL
Sbjct 540 LSEKERQIKKQALVELVKKPKATKEQLKAVMDDFAAFVEKCCAKDKETCFAEEGKKL 599
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Query 600 ISKTRAAALGV 609
++ ++AALG+
Sbjct 600 VAASQAALGL 609
XXXXXXXXXXXX
  
```

(a)

Fig. 2. Sequence alignment of HAFP (Query) to template molecules HSA (a) and VTDB (b) with indication of secondary structure elements. Here H — alpha-helix, X — irregular structure.

Score = 87.4 bits (215), Expect = 1e-21, Method: Compositional matrix adjust.  
 Identities = 101/485 (21%), Positives = 193/485 (40%), Gaps = 33/485 (7%)

Query	9	LIFLLNFTESTRTLHRN-EYGIASILDYQCTAEISLADLATIFFAQVQVEATYKEVSKMV	67
		L+ LL L R +Y + + + L+ + +++ T+++VS++V	
Sbjct	5	LVLLLAVAFGHALERGRDYEKKNVKCFESHGKEDFTSLSLVLYSRKFPSPGTFEQVSQLV	64
		XX	
Query	68	KDALTAIEKPTGDEQSSGCLLENQLPAFLEELCHEKEILEKY-GHSDCCSQSEEGRHNCFL	126
		K+ ++ E + C + + A + C + G ++CC++ R C	
Sbjct	65	KEVVSLEACCAEGADPCYDTRTSALSASCSNSPPFVHPGTAECCTKEGLERKLCMA	124
		XX	
Query	127	AHKKPTPASIPLFQVPEPVTSCAEYEEDRETFMNKFIYEIARRHPFLYAPTILL--WAAR	184
		A K P P + P CEA+ +D + + N+F++E + + AP LL +	
Sbjct	125	ALKH-QPQEFPTYVEPTNDEICEAFRKDPKEYANQFMWEYSTN--YGQAPLSLLVSYTKS	181
		HHXX-XX	
Query	185	YDKIIPSCCKAENAVECFQTKAATVTKELRESSLLNQHACAVMKNFGTRTFQAITVTKLS	244
		Y ++ SCC + + CF K K L + L+ C+ +G + + + KL+	
Sbjct	182	YLSMVGSCCTSASPTVCF-LKERLQLKHLSSLTTLNRCVQYAAAYGEKKSRLSNLIKLA	240
		XX	
Query	245	QKFTKVNFTIQLKLVLDVAHVHEHCCRGDVLDCD-LQDGEKIMSYICSQQDTLSNKITECC	303
		QK + ++ L D+ ++ CC DC+ ++ + +C T ++K +CC	
Sbjct	241	QKVPADLEDLVPLAEDITNILSKCCESASEDCMAKELPEHTVKLCDNLSKNSKFEDCC	300
		HHHXX	
Query	304	K-LTTLERGCQIHAENDEKPEGLSPNLNRFGLDRDFNQFSSGKNIIFLASFVHEYSRRH	362
		+ T ++ C P P L + + G + + + E SRR	
Sbjct	301	QEKMTAMDVFVCTYFM-----PAAQLPELDPVELPTNKDVCDPGNTKV-MDKYTFELSRRT	354
		XX	
Query	363	--PQLAVSVILR-VAKGYQELLE-----KCFQTENPLECQDKGEEELQKYIQESQALAKR	414
		P++ +S +L K E + CF + PL ++EL +I + Q L	
Sbjct	355	HLPEVFLSKVLEPTLKLGLGECDDVEDSTTCFNAKGPL-----LKKELSSFIDKGQEL---	406
		XX	
Query	415	SCGLFQKLGEYYLQNAFLVAYTKKAPQLTSSSELMAITRKMAATAATCCQLSEDKLLACGE	474
		C + + + K P T EL + K + A+ CC ++ L E	
Sbjct	407	-CADYSENTFTTEYKKLAERLKAQLPDATPKELAKLVNKRSDFASNCCSINSPLYCDSE	465
		-XX	
Query	475	GAADI 479	
		A++	
Sbjct	466	IDAEL 470	
		HHHHH	

(b)

Fig. 2. (Continued)

the template molecule (23%). Since variable regions and the majority of the sequence differences are typically localized in the loops, errors are often accumulated in loops. So, low amount of loop regions increases accuracy of the model.

The model generated was stored as PDB output file and then visualized and analyzed using UCSF Chimera package.<sup>47</sup> Visualization of the model showed that it has U-shaped structure with a cavity formed by domains I and III. So, the structure of our model is in accordance with experimental data obtained by combining dark-field

electron microscopy with laser-assisted optical system that allow producing relative mass map of a molecule.<sup>8</sup> This map showed asymmetrical structural features of HAFP molecule, notably, three-mass dense regions at both extremities and at the vertex of the molecule. Each of these regions corresponds to domains I, II and III of HAFP.

Figure 3 shows visualization of the overall HAFP 3D model in two types of representation. The first one is ribbon [Fig. 3(a)] representation with indication of three-domain structure and the second one is surface [Fig. 3(b)] representation where the cavity is well visualized. Walls in the cavity are predominantly formed by four  $\alpha$ -helices that contain, mainly, hydrophilic amino acid residues with hydrophobic areas that involve aa residues L138, V141, V145, L447, M448, A449, I450, M454, A455, A456, A458, A459. Hypothetical estrogen-binding site is located in this cavity and involves aa residues of the following alpha-helical regions: 144–158, 442–460, 466–490, 542–559 [Fig. 3(a)].

### 3.2. Validation of the model structure

Calculation of RMSD was performed for the model structure during relaxation by 5 ns equilibrium MD simulation in respect with the initial model structure. RMSD values (Fig. 4) for extended alpha-helical regions that form walls of the cavity that contain hypothetical estrogen-binding site of HAFP and encompass aa residues 144–158, 442–460, 466–490 and 542–599 are stabilized after 500 ps MD simulation

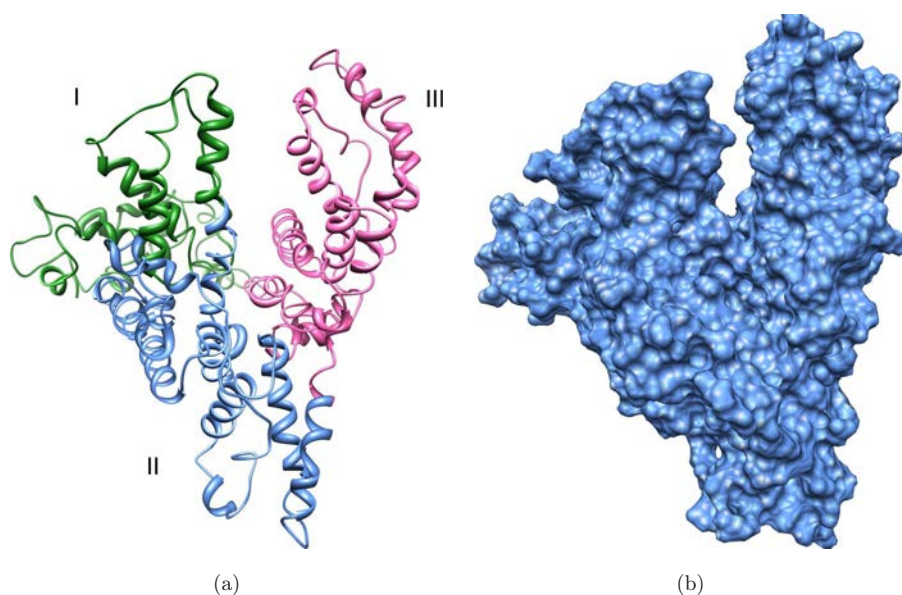


Fig. 3. Three-dimensional model of human AFP generated on the basis of homology with serum albumin and vitamin D-binding protein with the use of MODELLER software. (a) Ribbon representation shows secondary structure elements (alpha-helices). Domains I, II and III are indicated. (b) Representation of the surface of HAFP 3D model: U-shaped structure with the cavity.

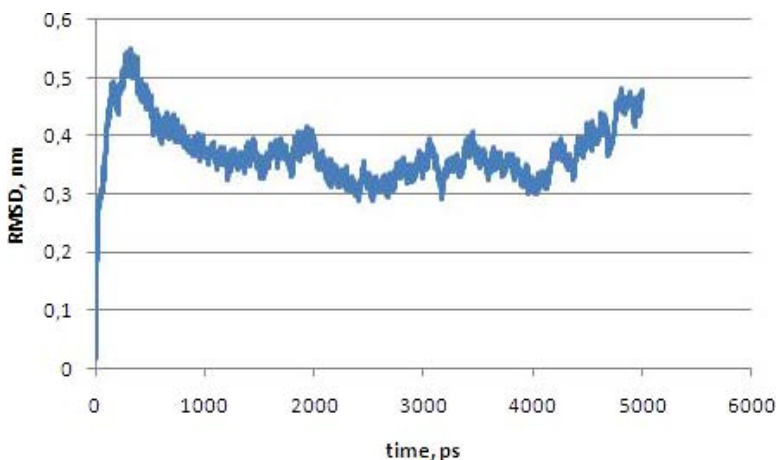


Fig. 4. RMSD plot received during 5 ns MD simulation for four extended alpha-helices located in the cavity of HAFP molecule where putative estrogen-binding site is localized. RMSD values for estrogen binding site demonstrate metastability of its structure.

and vary between 0.30 and 0.40 nm indicating metastability of the binding site conformation. Variations in RMSD value were due to internal dynamics of the protein and low frequency fluctuations and motions of four alpha helices in relation to each other in explicit water environment.

PROCHECK (version 3.5) program was further used for the analysis of the model quality. To estimate conformational properties of all aa residues of the 3D model, the program draws Ramachandran map that contains values of combinations of the protein backbone dihedral angles  $\varphi$  and  $\psi$  for its each residue. Analysis of the Ramachandran map for amino acid residues of the obtained HAFP 3D model showed high-quality parameters for conformations of all non-glycine and non-proline residues (540 in total). The plot statistics was as follows. The main chain conformations for 498 (92.2%) of these amino acid residues were within the most favored regions, 41 (7.6%) residues were in the additional allowed regions, 0% — in the generously allowed regions and only 1 (0.2%) residue (A39) was located in disallowed region (Fig. 5). The number of glycine and proline residues were 26 and 21, respectively.

Individual residue geometrical properties may also be used to assess the quality of the model because they may vary along a protein sequence. PROCHECK program draws various graphs and diagrams that give information about regions with consistently poor or unusual geometry. These diagrams drawn for our model showed that absolute deviations from mean values of torsion angles  $\chi_1$ ,  $\omega$  and  $\zeta$  ( $C\alpha$ -chirality) are low enough (Fig. 6). The amount of residues that deviate by more than 2.0 standard deviations from ideal, i.e. with unusual geometry (R121, F202, Q291, K304, K384, C500, F533, T564 and E566) was small and consisted only 1.53% from the total amount of amino acid residues.

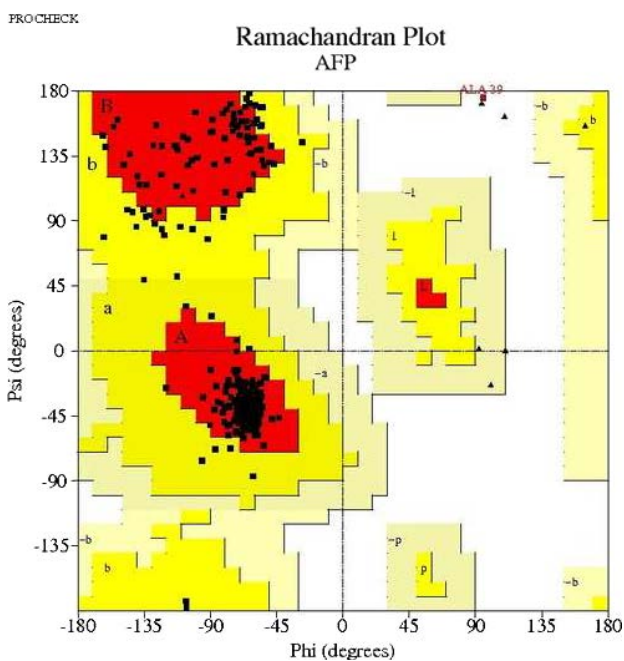


Fig. 5. Ramachandran map for the main chain conformations of all residues of HAFP model drawn by PROCHECK program. Combinations of  $\varphi$  and  $\psi$  torsion angles demonstrate that all non-glycine and non-proline residues (except A39) are in the most favored and allowed regions.

Secondary structure elements correspond to  $\alpha$ -helices and random coils, with no  $\beta$ -structures. The amount of  $\alpha$ -helices is in accordance with experimental data and constitutes about 69% from the modeled polypeptide chain length. Experimentally obtained data show that amount of  $\alpha$ -helical regions comprise up to 67% from full-length molecule. Our model has a little higher amount of  $\alpha$ -helices than in experiments because of absence of  $N$ -terminal part of the polypeptide chain that seems to be predominantly unstructured in the native molecule.

Estimation of water accessibility areas showed that four  $\alpha$ -helices that encompass aa residues (142–158, 442–460, 466–490 and 542–559) of putative estrogen-binding site are mostly buried indicating enrichment in hydrophobic aa residues. The program estimated also  $G$ -factor values for combinations of  $\varphi$ - $\psi$ ,  $\chi_1$ - $\chi_2$ ,  $\chi_3$ - $\chi_4$ , and  $\omega$  torsion angles along with the main chain bonds and angles. Analysis of all these diagrams also showed that the model constructed is reliable and useful for further characterization of the protein.

### 3.3. Docking DES molecule to HAFP model

Low RMSD values calculated for hypothetical estrogen-binding site indicate reasonable good quality of this part of the model and allow performing of docking of the ligand to HAFP model constructed.

## Residue properties AFP

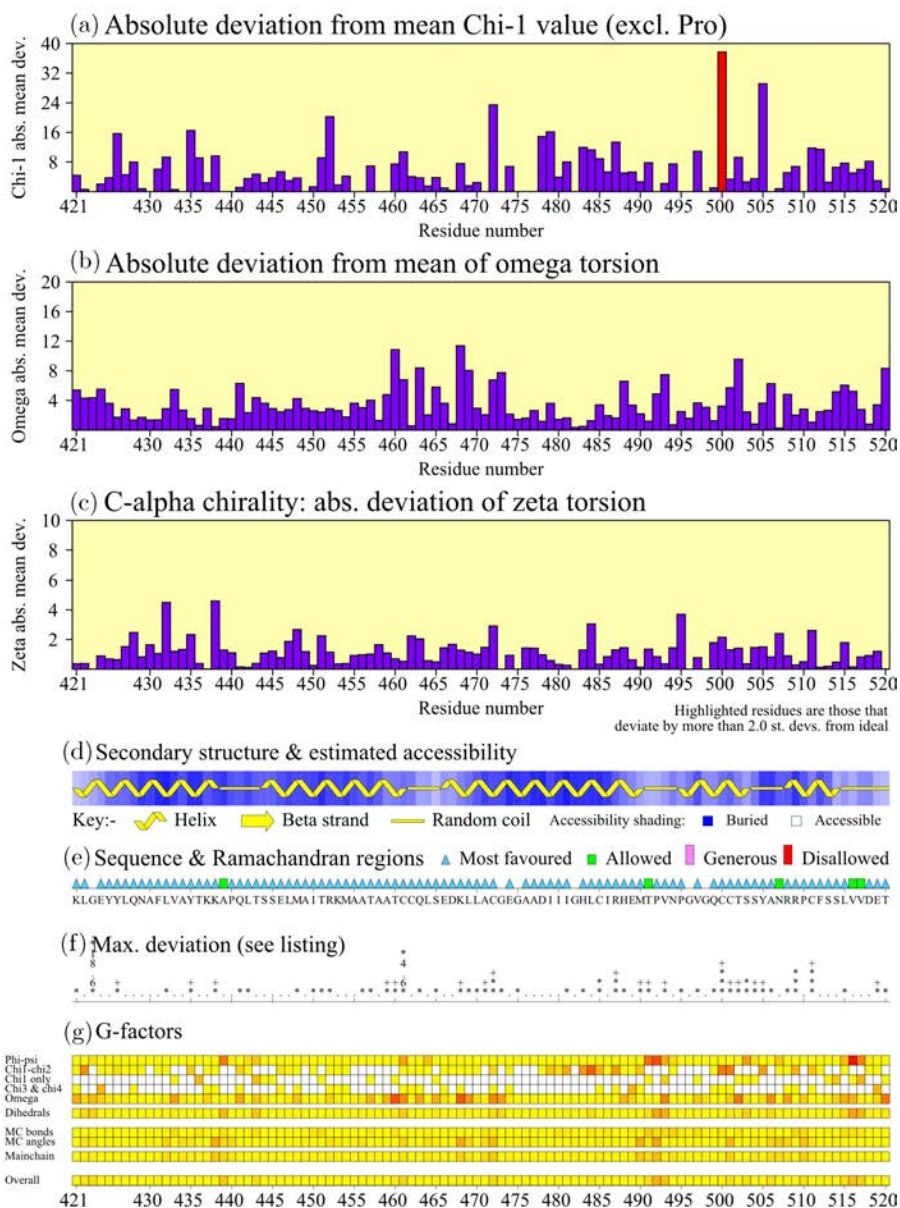


Fig. 6. Stereochemical properties of individual aa residues of HAFFP 3D model. Here the part of polypeptide chain that contains amino acid residues of putative estrogen-binding site is shown. The parameters shown in the form of graphs and diagrams indicate high quality of the model.

The AutoDock (version 4.0) software package was used to perform partly flexible (induced fit) molecular docking with the aim to search global energy minimum of interaction between the HAFP model and DES. In this way, the docking studies allow exploring the real binding environment, where the ligand molecule interacts with the target protein. This is especially valuable in case if a structure of ligand-binding site is not experimentally obtained.

In this study, the following aa residues were given as flexible: Q428, M448, R452, V543 and T547. It was shown that DES molecule is docked to a part of an extended  $\alpha$ -helical structure on the bottom of the cavity in HAFP model, including cluster of hydrophobic aa residues (Fig. 7). Aromatic benzene rings of DES, seemingly, interact with hydrophobic aa residues M448, A449 and V543. The interaction is also provided by hydrogen bonding: oxygen atoms in the side chains of residues S445 and E551 are seemed to form hydrogen bonds with hydroxyl groups of DES. The corresponding distances are: 1.95 Å between OG atom of residue S445 and H-atom of the first hydroxyl (HO1) of DES, and 1.69 Å between OE1 atom of residue E551 and the H-atom of the second hydroxyl group (HO2) of the hormone.

Interestingly, all interacting aa residues belong to subdomains IIIA and IIIB of HAFP. The thermodynamic characteristics of DES binding to HAFP was estimated using value of free energy change ( $\Delta G$ ) that was equal to  $-4.24$  kcal/mol for the best docked complex.

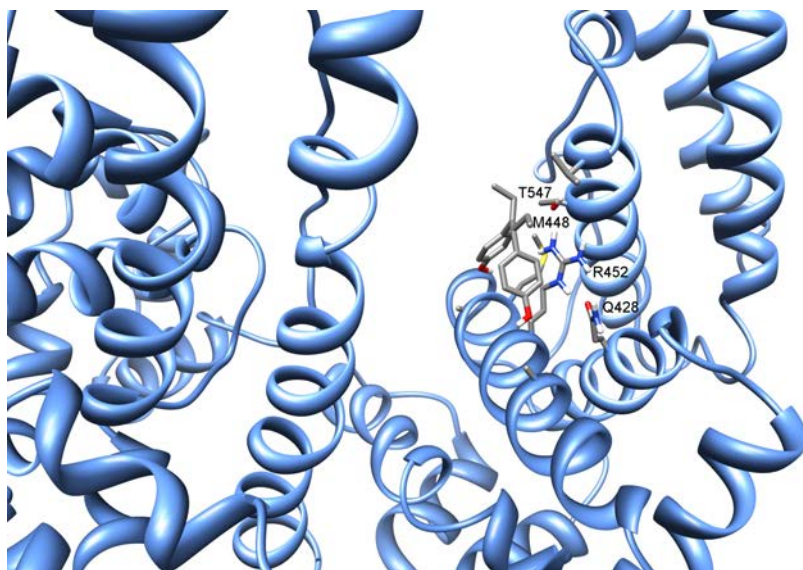


Fig. 7. Docking of DES to estrogen-binding site of the 3D model of HAFP. Amino acid residues that are localized close to the hormone are shown. DES is docked to alpha-helices on the bottom of the cavity. Free energy change ( $\Delta G$ ) was estimated as  $-4.24$  kcal/mol for the best complex.

### 3.4. Putative estrogen-binding residues revealed by MD simulation

Molecular dynamics simulation of the HAFP-DES complex was performed based on the docked complex structure to obtain a dynamical picture of conformational changes along with detailed and precise information about interaction between HAFP and the hormone molecule. The trajectory from the 5 ns MD simulation was used to extract data about types of interaction and aa residues that participate in them. Visualization of the complex (Fig. 8) shows that the hormone molecule changes its position during MD simulation and interacts not only with aa residues of the main  $\alpha$ -helices in subdomains IIIA and IIIB (S445, M448, R452, M548 and E551), but also aa residues located in subdomain IB (L138).

To calculate distances for hydrophobic interactions, we took centers of mass of side chains of residues L138, M448, I450 and M548 and aromatic rings of the hormone. This increases interaction distances because the van der Waals radii of all atoms of side chains of the residues should be taken into account. Here, the smallest distances are observed between aromatic ring A of DES and centers of masses of the residues L138 and M448 side chains: they vary, mainly, from 0.50 to 0.80 nm [Figs. 9(a) and 9(b)].

For hydrogen bonding, distances were calculated between separate atoms. The smallest distances were observed between S445, E551 and second hydroxyl group of DES and also between R452 and first hydroxyl group of DES. Figure 9 depicts distances between two pairs of oxygen atoms: between OE1 atom of glutamic acid E551 and OH2 atom of the second hydroxyl group of DES [Fig. 9(c)] and between OG atom of serine S445 and OH2 atom of DES [Fig. 9(d)]. Small distances observed after 2 ns of MD simulation indicate possibility of interaction between these atoms through formation of hydrogen bonds.

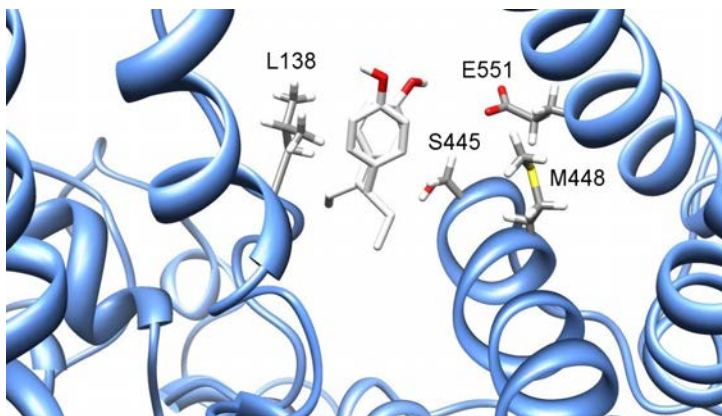
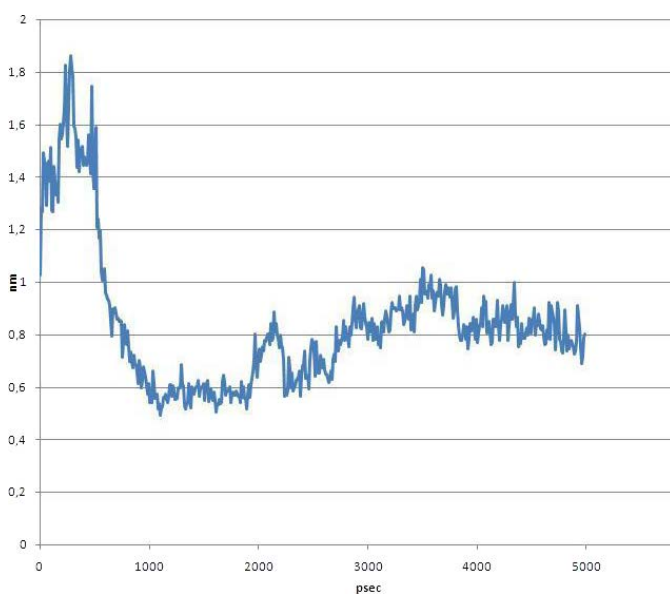
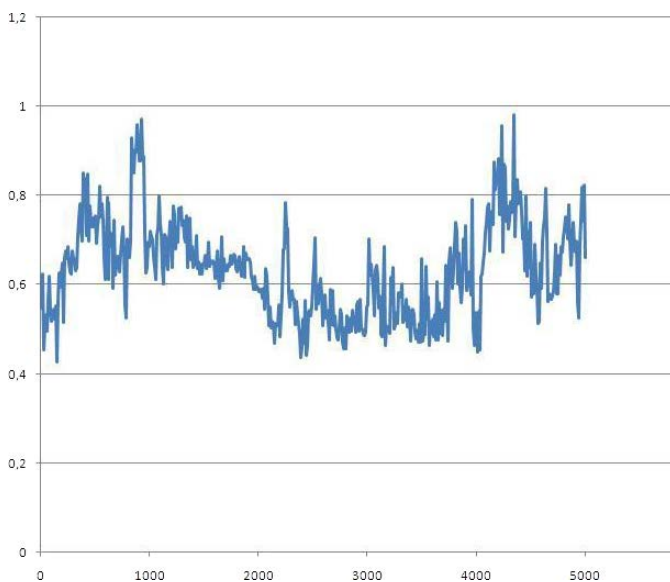


Fig. 8. Formation of the complex between HAFP 3D model structure and DES, revealed by molecular dynamics (MD) simulation study. Amino acid residues involved in interaction between HAFP and DES are shown.



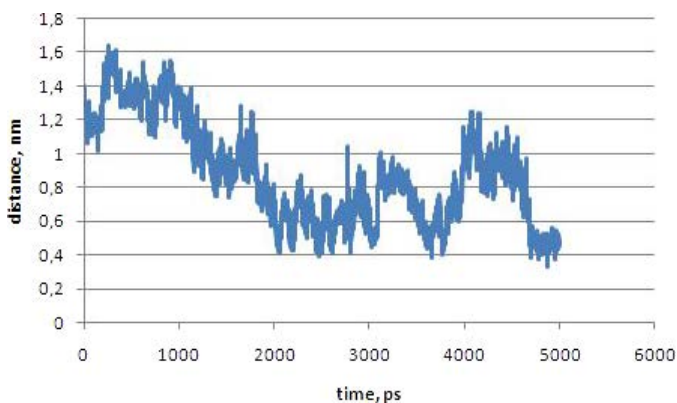


(a)

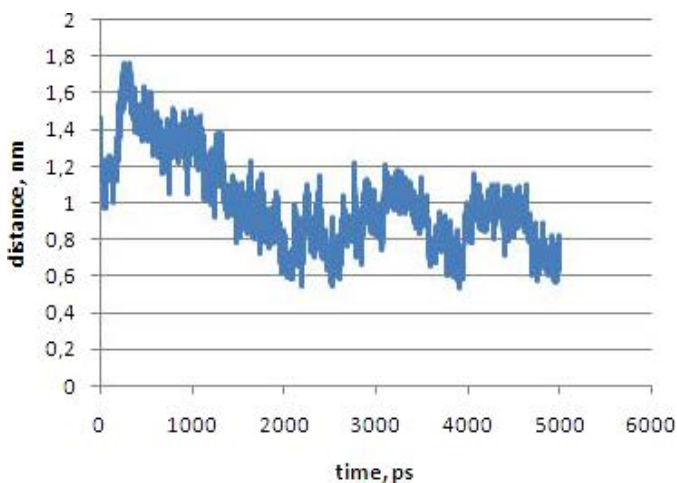


(b)

Fig. 9. Time-dependent changes in distances between amino acid residues of HAFP 3D model and functional groups of DES during 5 ns MD simulation. For van der Waals interactions changes in distances between centers of masses of hydrophobic amino acid L138 (a) and M448 (b) side chains and benzene rings of DES are shown. For hydrogen bonding changes in distances between polar atoms are shown: OE1 atom of glutamic acid E551 and OH2 atom of the second hydroxyl group of DES (c); and OG of serine S445 and OH2 atom of DES (d). Small distances observed after 2 ns of MD simulation indicate possibility of interaction between these atoms and of hydrogen bonding.



(c)



(d)

Fig. 9. (Continued)

A conclusion may be made that interactions between the hormone and the estrogen-binding site is, seemingly, provided by two sets of amino acid residues: (1) hydrophobic side chains of amino acid residues in HAFP (L138, M448 and M548) that interact with aromatic rings of DES; and (2) hydrogen bonding with participation of OH-groups of the hormone and polar atoms in HAFP amino acid side chains (S445, R452 and E551).

#### 4. Discussion

Human AFP domains I and II include 192 amino acids each (residues 19–210 and 211–402, respectively); domain III consists of 199 residues (residues 403–601). The

highest identity degree between primary structures of HAFP and HSA is found in the domain III (48%), whereas the lowest identity is found in the subdomain IA (16%).<sup>48</sup> Since the method of homology modeling relies on the assembly of conserved structural elements and also because variable regions are typically located in the loops where the majority of the sequence differences are localized, it was necessary to estimate the secondary structure elements. Pairwise alignment with indication of secondary structure elements in both the template molecules shows (Fig. 2) that amount of flexible loop regions is not high (about 23% in HSA and 22% in VTDB). Pairwise comparison of primary structures of AFP and albumin from various animal species revealed that identity degree between these proteins reduced from human to dog, horse, mouse, and rat. So, a divergence vector from a common ancestor increases in this order and demonstrates the highest conservation of aa sequences for the pair of HAFP–HSA. All these observations along with the fact that 3D structures of proteins are more conserved than their primary structures allow us to construct 3D model of HAFP on the basis of homology with HSA.

Validation of the entire model of HAFP using PROCHECK v3.5 program showed its good quality and reliability. Indeed, the Ramachandran plot drawn for the model structure indicates that the amount of amino acid residues with conformations in the most favored region is more than 90% (92.2%) and the total amount of residues in the most favored and allowed regions is 99.8%. An assessment of residue-by-residue stereo-chemical parameters showed that amino acid residues from 446 to 490 that compose the proposed estrogen-binding site have no unusual stereo-chemical properties.

The model constructed in this work has a U-shaped structure (Fig. 3) and this shape corresponds to that obtained experimentally by methods of electron microscopy, image processing and circular dichroism.<sup>8</sup> The overall architecture of the model corresponds to three-domain structure demonstrated by these methods. The secondary structure elements of the model also correspond to those obtained in experiments that demonstrated existence of  $\alpha$ -helices and random coils with no  $\beta$ -structures. Also, the amount of  $\alpha$ -helical regions is close to the experimental data. So, the model structure is in good agreement with the data obtained experimentally, and this proves accuracy of the model generated in this work.

The model obtained explains that the U-shaped structure of HAFP provides formation of the cavity between domains I and III. It is thought that this shape of the entire molecule along with existence of the cavity is not casual. This is of special importance since the putative estrogen-binding sites (both high and low affinity) are located in this cavity. Earlier, with the use of affinity chromatography method, we showed that both free and immobilized diethylstilbestrol has the highest efficiency of binding to HAFP in comparison to natural estrogens.<sup>22</sup> As for naturally occurring estrogens (17 $\beta$ -estradiol and estriol), it has been shown that only immobilized hormones are able to bind to HAFP, suggesting the importance of spatial orientation of the hormone molecules to the estrogen-binding site in *in vivo* conditions. Probably, it is because of the planarity of structures of naturally occurring estrogens. However,

binding to HAFP of both natural estrogens and their synthetic analog (DES) depends on the length of spacer and the mode of immobilization. It may be suggested that the immobilization of the hormone using a long spacer provides access of estrogen molecule to the binding site and, thereby, facilitates interaction between HAFP and estrogens.

Also, it may be proposed that the conformational changes that take place in HAFP molecule are of crucial importance during estrogen binding. It has been demonstrated that high concentrations of estrogens (in ratio of molar concentrations of AFP to E2 as 1:1280) cause conformational changes in the AFP molecule.<sup>26</sup> It has also been shown that domains I and III of HAFP have a rigid tightly packed tertiary structure, whereas domain II has the conformation that corresponds to the molten globule form (MGF). The C-terminal part of domain II is considered as a “hinge region,” responsible for conformational flexibility of all domains.<sup>8</sup> This promotes interaction of ligand molecules, including estrogens with their binding sites in HAFP molecule.

The calculation of RMSD values was performed for four extended  $\alpha$ -helices located in the cavity that contains a hypothetical estrogen-binding site and was of special interest to us. This method showed metastability and reasonable good quality of this part of the model. This allows docking of DES molecule to the model structure with subsequent MD simulation study that revealed amino acid residues involved in the hormone-binding. On the basis of data obtained it is suggested that interaction between DES and HAFP takes place through: (1) formation of hydrogen bonds with participation of OH-groups of the hormone and polar atoms of S445, R452, and E551 residues; and (2) van der Waals interactions with involvement of hydrophobic side chains of L138 and M448 residues and aromatic benzene rings (and ethyl groups) of DES. It may be concluded that most aa residues that bind DES are located in the subdomains IIIA and IIIB of HAFP. Importantly, these data are in agreement with experimentally obtained data, which demonstrated that estrogen-binding amino acid residues should be located in the domain III of HAFP.<sup>23–26</sup>

The model constructed in this work should also be validated in experiments. This can be done, for example, by rational design of new ligands and their subsequent testing *in vitro* or *in vivo*.

## References

1. Deutsch HF, Chemistry and biology of alpha-fetoprotein, *Adv Cancer Res* **56**:253–312, 1991.
2. Terentiev AA, Moldogazieva NT, Structural and functional mapping of alpha-fetoprotein, *Biochemistry (Mosc.)* **71**:120–132, 2006.
3. Mizejewski GJ, Biological role of alpha-fetoprotein in cancer: Prospects for anticancer therapy, *Expert Rev Anticancer Ther* **2**:709–735, 2002.
4. Morinaga T, Sakai M, Wegmann TG, Tamaoki T, Primary structure of human alpha-fetoprotein and its mRNA, *Proc Natl Acad Sci USA* **80**:4604–4608, 1983.
5. Pucci P, Siliciano R, Malorni A, Marino G, Tecce MF, Ceccarini C, Terrana B, Human alpha-fetoprotein primary structure: A mass spectrometric study, *Biochemistry* **30**:5061–5066, 1991.

6. Song Y-H, Naumova AK, Liebhaber SA, Cooke NE, Physical and meiotic mapping of the region of human chromosome 4q11-q13 encompassing the Vitamin D-binding protein DBP/Gc-globulin and albumin multigene cluster, *Genome Res* **9**:581–587, 1999.
7. Beatie WG, Dugaiczky A, Structure and evolution of human alpha-fetoprotein deduced from partial sequence of cloned cDNA, *Gene* **20**:415–422, 1982.
8. Luft AJ, Loscheider FL, Structural analysis of human and bovine  $\alpha$ -fetoprotein by electron microscopy, image processing and circular dichroism, *Biochemistry* **22**:5971–5978, 1983.
9. He XM, Carter DC, Atomic structure and chemistry of human serum albumin, *Nature* **358**:209–215, 1992.
10. Petitpas I, Gruene T, Bhattacharya AA, Curry S, Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids, *J Mol Biol* **314**:955–960, 2001.
11. Zunszain PA, Ghuman J, Komatsu T, Tsuchida E, Curry S, Crystal structure analysis of human serum albumin complexed with hemin and fatty acids, *BMC Struct Biol* **3**:6, 2003.
12. Petitpas I, Petersen CE, Ha CE, Bhattacharya AA, Zunszain PA, Ghuman J, Bhagavan NV, Curry S, Structural basis of albumin-thyroxine interactions and familial dysalbuminemic hyperthyroxinemia, *Proc Natl Acad Sci USA* **100**:6440–6445, 2003.
13. Ghuman J, Zunszain PA, Petitpas I, Bhattacharya AA, Otagiri M, Curry S, Structural basis of the drug-binding specificity of human serum albumin, *J Mol Biol* **353**:38–52, 2005.
14. Zunszain PA, Ghuman J, McDonagh AF, Curry S, Crystallographic analysis of human serum albumin complexed with 4Z,15E-bilirubin-IXalpha, *J Mol Biol* **381**:394–406, 2008.
15. Verboven C, Rabijns, A, De Maeyer M, Van Baelen H, Bouillon R, De Ranter C, A structural basis for the unique binding features of the human Vitamin D-binding protein, *Nat Struct Biol* **9**:131–136, 2002.
16. Otterbein LR, Cosio C, Graceffa P, Dominguez R, Crystal structures of the vitamin D-binding protein and its complex with actin: Structural basis of the actin-scavenger system, *Proc Natl Acad Sci USA* **99**:8003–8008, 2002.
17. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE, The protein data bank, *Nucleic Acids Res* **28**:235–242, 2000.
18. Chothia C, Lesk AM, The relation between the divergence of sequence and structure in proteins, *EMBO J* **5**:823–826, 1986.
19. Kaczanowski S, Zielenkiewicz P, Why similar protein sequences encode similar three-dimensional structures? *Theor Chem Accoun* **125**:543–550, 2010.
20. Hsia JC, Er JS, Tan CT, Ester T, Ruoslahti E,  $\alpha$ -Fetoprotein binding specificity for arachidonate, bilirubin, docosahexaenoate and palmitate, *J Biol Chem* **255**:4224–4227, 1980.
21. Uriel J, de Nechaut B, Dupiers M, Estrogen-binding properties of rat, mouse and man fetospecific serum protein. Demonstration by immuno-autoradiographic methods, *Biochem Biophys Res Commun* **46**:1175–1180, 1972.
22. Tatarinov YuS, Terentiev AA, Moldogazieva NT, Tagirova AK, Human alpha-fetoprotein and its purification by chromatography on immobilized estrogens, *Tumor Biol* **12**:125–130, 1991.
23. Shahbazzadeh D, Estrogen binding activities of recombinant alpha-fetoproteins expressed in yeast, *Hokkaido Igaku Zasshi* **70**:473–483, 1995.
24. Nishi S, Matsue H, Yoshida H, Yamamoto R, Sakai M, Localization of estrogen-binding site of alpha-fetoprotein in chimeric human-rat proteins, *Proc Natl Acad Sci USA* **88**:3102–3105, 1991.
25. Mizejewski GJ, Biological roles of alpha-fetoprotein during pregnancy and prenatal development, *Exp Biol Med* **229**:439–463, 2004.

26. Vakharia D, Mizejewski GJ, Human alpha-fetoprotein peptides bind estrogen receptor and estradiol and suppress breast cancer, *Breast Cancer Res Treat* **63**:41–52, 2000.
27. Venclovas C, Margelevicius M, Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment, *Proteins* **61**(S7):99–105, 2005.
28. Boeckmann B, Blatter M-C, Famiglietti L, Hinz U, Lane L, Roechert B, Bairoch A, Protein variety and functional diversity: Swiss-Prot annotation in its biological context, *Comptes Rendus Biologies* **328**:882–899, 2005.
29. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, Clustal W and Clustal X version 2.0, *Bioinformatics* **23**:2947–2948, 2007.
30. Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A, Comparative Protein Structure Modeling With MODELLER, *Curr Protocol Bioinform Suppl* **15** 5.6.1–5.6.30, 2006.
31. Zhang Y, Progress and challenges in protein structure prediction, *Curr Opin Struct Biol* **18**:342–348, 2008.
32. Ginalski K, Comparative modeling for protein structure prediction, *Curr Opin Struct Biol* **16**:172–177, 2006.
33. Zhang Y, Skolnick J, The protein structure prediction problem could be solved using the current PDB library, *Proc Natl Acad Sci USA* **102**:1029–1034, 2005.
34. Maiorov VN, Crippen GM, Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins, *J Mol Biol* **235**:625–634, 1994.
35. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM, Stereochemical quality of proteins structure coordinates, *Proteins* **12**:345–364, 1992.
36. Laskowski RA, MacArthur MW, Moss DS Thornton JM, PROCHECK: A program to check the stereochemical quality of protein structures, *J Appl Cryst* **26**:283–291, 1993.
37. Froimowitz M, HyperChem: A software package for computational chemistry and molecular modeling, *Biotechniques* **14**:1010–1013, 1993.
38. Pearlman DA, Case DA, Caldwell JW, Seibel GL, Singh UC, Weiner P, Kollman PA *Amber 4.0*, University of California, San Francisco, 1991.
39. Yang L, Tan CH, Hsieh MJ, Wang J, Duan Y, Cieplak P, Caldwell J, Kollman PA, Luo R, New-generation Amber united-atom force field, *J Phys Chem B* **110**:13166–13176, 2006.
40. Schmidt MW, Baldrige KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA, General atomic and molecular electronic structure system, *J Comput Chem* **14**:1347–1363, 1993.
41. Mulliken RS, Electronic population analysis on LCAO-MO molecular wave functions, *I J Chem Phys* **23**:1833–1831, 1995.
42. Hinchliffe A, *Modelling Molecular Structures*, 2nd ed., England: John Wiley and Sons Ltd. p. 186, 2000.
43. Baker D, Sali A, Protein structure prediction and structural genomics, *Science* **294**:93–96, 2001.
44. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ, Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function, *J Comput Chem* **19**:1639–1662, 1998.
45. Berendsen HJC, van der Spoel D, van Drunen R, Gromacs — A message-passing parallel molecular-dynamics implementation, *Comp Phys Comm* **91**:43–56, 1995.
46. Gibbs PEM, Witke WF, Dugaiczky AJ, The molecular clock runs at different rates among a closely related members of a gene family, *Mol Evol* **46**:552–561, 1998.

47. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE, UCSF Chimera — A visualization system for exploratory research and analysis, *J Comput Chem* **25**:1605–1612, 2005.
48. Barker ME, Evolution of alpha-fetoprotein: Sequence comparisons among AFP species and with albumin species, *Tumor Biol* **9**:123–136, 1988.



**Alexander Alexandrovich Terentiev** is Professor and Head of Biochemistry Department at Russian State Medical University (RSMU), Moscow, Russia. He received his Ph.D. degree in 1969 and his D.Sci degree in 1990 from RSMU. His fields of scientific interest include chemistry and immunochemistry of oncofetal proteins, physicochemical properties and structure/function relationships in proteins, searching of biologically active sites by bioinformatics tools. He is author of more than 100 scientific publications and 33 inventions. He has been invited as a referee and reviewer for a number of scientific journals such as *Clin Chim Acta*, *Atlas Genet Cytogenet Oncol Hematol*, etc.



**Nurbubu Tentievna Moldogazieva** is an Associate Professor of Biochemistry Department at Russian State Medical University (RSMU), Moscow, Russia. She received her Ph.D. degree in 1988 from RSMU. Her thesis was devoted to study of estrogen-binding capacity of human and rodent alpha-fetoproteins. She is currently engaged in bioengineering and bioinformatics, computational biology, study of structure/function relationships in proteins, molecular dynamics and computer-based molecular modeling. She is author of 88 scientific publications. She has been invited as a referee and reviewer for scientific journals and projects such as *J Peptide Science*, *INTAS*, etc.



**Olga Vladimirovna Levtsova** graduated from M.V. Lomonosov Moscow State University (MSU), Biological Faculty, Department of Biophysics in 2005. She has Ph.D. degree in physics and mathematics (received in 2008 from MSU). Her thesis had interdisciplinary (biology, physics and mathematics) character and was devoted to molecular modeling of antimicrobial peptides and their interaction with membranes. Now she is working on computer-based modeling of interaction between membrane, receptors and small ligands, molecular docking and molecular dynamics. She is author of 25 scientific publications.



**Dmitry Maximenko** is a graduate student at Russian State Medical University (RSMU), Moscow, Russia. He started his scientific work in 2009 and is currently interested in the field of computer-based modeling of proteins and peptides, molecular docking and molecular dynamics. He participates in and makes reports on scientific conferences in the field of computational biology and bioinformatics.



**Denis Borozdenko** is a graduate student at Russian State Medical University (RSMU), Moscow, Russia. He started his scientific work in 2010 and is currently engaged in research in the field of computer-based modeling of proteins and peptides, molecular docking and molecular dynamics. He participates in and makes reports on scientific conferences in the field of computational biology and bioinformatics.



**Konstantin Voldemarovich Shaitan** is Professor in Biophysics and Deputy Director of Department of Bioengineering of Biological Faculty at M.V. Lomonosov Moscow State University (MSU), Russia. He received his Ph.D. degree in 1974, and D.Sci. degree in physics and mathematics in 1985 from MSU. His scientific fields of interest are in biophysics and computational biology. He is engaged in study of biomolecules by molecular dynamics and in development of algorithms and software. He is author of about 300 scientific publications. He has been invited as a referee and reviewer for a number of scientific journals, projects and programs.