Theor Chem Acc (1999) 101:51–56
DOI 10.1007/s002149800m84

Theoretical
Chemistry Accounts

Regular article

Molecular modeling of the hypoxia-inducible factor-1 (HIF-1)*

G. Michel[1], E. Minet[2], I. Ernest[2], F. Durant[1], J. Remacle[2], C. Michiels[2]

[1]Laboratoire de Chimie Moléculaire Structurale, Facultés Universitaires Notre-Dame de la Paix, rue de Bruxelles 61, B-5000 Namur, Belgium
[2]Laboratoire de Biochimie et Biologie Cellulaire, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium

Received: 24 April 1998 / Accepted: 4 August 1998 / Published online: 11 November 1998

Abstract. Hypoxia-inducible factor-1 (HIF-1) is a heterodimeric transcription factor activated by hypoxia. It is composed of two different subunits, HIF-1α and ARNT (aryl receptor nuclear translocator). When activated, HIF-1 mediates the differential expression of genes such as erythropoietin and vascular endothelial growth factor (VEGF). This work is aimed at defining the 3D structure of HIF-1. In this work we use the powerful modeling approach in order to obtain a first model of HIF-1. The two subunits HIF-1α and ARNT belong to the bHLH (basic helix-loop-helix) PAS (per, ahr/arnt, sim) family. We focused on the bHLH domain since the protein/DNA interactions are mediated by the basic domains and the dimerization by hydrophobic interactions between the two helices. Firstly, we aligned the sequence of each subunit along with sequences of other bHLH factors, using the program CLUSTALW. Such alignments could not detect the sequence corresponding to the bHLH domain. We therefore used the program GIBBS, which detects common motifs between sequences of different proteins. We indeed obtained common residues which could constitute the bHLH motif. In order to validate the target sequences, we submitted both sequences to a secondary structure prediction algorithm, PHD. Secondly, we made sequence/structure alignments in order to find a template using the BLAST program. We were then able to obtain a structural model of the heterodimer by means of HOMOLOGY. Once the four-helix bundle was constructed, we generated both loops using the program SCWRL for adding sidechains to the protein backbone based on the backbone-dependent library. Thirdly, we minimized the model of the heterodimer to avoid steric clashes in its structure, using the program DISCOVER. A combination of tools, such as Ramachandran plots, physicochemical properties, and energetic profiles, enabled us to validate the 3D model. The next step will be to dock the structure onto the DNA recognition site. This first study already allowed us to obtain a structural view of HIF-1. It has to be confirmed by X-ray structure analysis of the protein and we will then be able to understand the protein DNA interaction of this transcription factor and its role played in the cellular response to hypoxia.

Key words: Hypoxia-inducible factor-1 (HIF-1) – Transcription factor bHLH – Protein-protein interactions – Hypoxia

1 Introduction

Hypoxia-inducible factor-1 (HIF-1) is a heterodimeric transcription factor activated by hypoxia. It is composed of two different subunits, HIF-1α and ARNT (aryl receptor nuclear translocator). The two subunits HIF-1α and ARNT belong to the bHLH (basic helix-loop-helix) PAS (per, ahr/arnt, sim) family, and are respectively composed of 826 and 789 amino acids [1].

HIF-1α is activated by hypoxia, migrates to the nucleus, and dimerizes with ARNT to form the active transcription factor HIF-1. HIF-1 recognizes the hypoxia-response element (HRE) [5′-(G/C/T)-ACGTGC-(G/T)-3′] present in the enhancers of several genes and leads to their overexpression [2, 3]. Amongst these genes are erythropoietin, vascular endothelial growth factor (VEGF), and several glycolytic enzymes. Erythropoietin is the glycoprotein hormone which regulates mammalian erythrocyte production and, as a result, oxygen delivery to tissue [3]. VEGF has been suggested to be a key mediator in hypoxia-induced angiogenesis [4]. HIF-1 is present in various cell types and seems to be part of a widespread O_2-sensing and signal transducing mechanism. It thus seems to play an important role in the adaptive response of tissue to hypoxia [5, 6].

This paper describes the results obtained for molecular modeling of HIF-1. We focused our work on the bHLH domain, which is composed of two amphipatic α-helices connected by a loop. This domain is responsible

*Contribution to the Proceedings of Computational Chemistry and the Living World, April 20–24, 1998, Chambery, France

Correspondence to: G. Michel
e-mail: gaetan.michel@fundp.ac.be
Tel.: +32-81-724569, Fax: +32-81-724530

for dimerization and binding to the consensus sequence of DNA.

2 Methods

2.1 Identification of the bHLH motif

2.1.1 Multiple alignments

Alignments are done using CLUSTALW [7]. The multiple alignments are carried out in three stages: (1) all sequences are compared to each other (pairwise alignments); (2) a dendrogram (like a phylogenetic tree) is constructed, describing the approximate groupings of the sequences by similarity; (3) the final multiple alignment is carried out, using the dendrogram as a guide. For the multiple alignment the selected matrix was BLOSUM62, which describes how similar are the sequences to be aligned. As an input, we used the sequence portions of four bHLH transcription factors and the entire sequences of HIF-1α and ARNT.

2.1.2 Motif detection

The software GIBBS [8] is based on three fundamental characteristics. First, the algorithm seeks a small number of sequence patterns throughout the sequences previously used as input. Second, a single pattern is described by a probabilistic model of residue frequencies at each position. Third, the location pattern within the target sequence is described by a set of probabilistically inferred position variables. As an input, we use 25 sequences of bHLH transcription factors, including the entire sequences of ARNT and HIF-1α [8].

The secondary structure predictions are predicted by the PHD [9, 10] algorithm. A sequence database is scanned for similar sequences to the user's sequence. Then, a multiple sequence alignment is generated by a weighted dynamic programming method and motifs are retrieved from the PROSITE database in order to compare the target protein to a domain database (ProDom). Finally, the multiple alignment is used as input for profile-based neural network predictions [9, 10].

2.2 Generation of a 3D model

The algorithm used to detect a template is BLAST (basic local alignment search tool) [11]. BLAST firtsly looks up an "index" of every oligomer in the database for oligomers showing a sufficient degree of similarity to those present in the query sequence. The program then tries to extend the initial regions of similarity into a larger ungapped alignment. A score taken from a scoring matrix (BLOSUM 62) is given to each amino acid pair and the alignment score is calculated by adding up individual amino acid pair scores. The database used was the Brookhaven protein data bank (PDB) [12] since the purpose of this alignment was to detect a 3D canvas. The query sequences corresponded to the "consensus" sequences defined in the previous part.

In the module HOMOLOGY [13] we assign the 3D coordinates of the template's residues to the residues corresponding to the helical structures in the "consensus" sequence. The loops are built using the Generate loop command in the HOMOLOGY module. This last command builds a peptide backbone chain between two anchor residues using randomly generated values for all the loops' phi and psi angles [13]. The loops are screened and those with unacceptable contacts are rejected. However, the conformations for the side chains are fully extended. We then used the program SCWRL [14] to position the sidechains in an ideal conformation. SCWRL is a program designed for adding sidechains to a protein backbone based on a backbone rotamer library. The program starts with the mainchain atoms (N, Cα, C, O) from the protein structure and then the calculated dihedral angles are used to select a list of rotamers from a rotamer library, for each residue. The rotamers are built, based on the parameter set from AMBER 4.1, and the set is searched for the minimum steric clashes to build the output structure. The structures obtained were optimized by molecular mechanics, using a three-steps process (steepest descent, conjugate gradient, VA09A) in the cuff forcefield with DISCOVER [15].

2.3 Analysis of the structure

Different tools were used to analyze the structure. Firstly, we calculated physicochemical properties using the program HINT [16]. HINT (Hydropathic INTeractions) is designed to detect and estimate the strength of the hydrophobic and polar interactions. The solvent partition calculations performed by HINT are based on the Hansch and Leo [17] approach to the hydrophobic fragment constant. The program then assigns hydrophobic constants and factors to each atom in the molecule. We chose a level of -20 to represent the polar contributions and a level of 5 for the hydrophobic contributions of the residues.

Secondly, we used PROSA [18] to determine errors in protein 3D structures. This method is based on the Boltzmann's principle. Forces are extracted from a database of known structures in the form of a potential of mean force [19] which constitutes the mean field. In a given structure the interaction energy e_{ij} between amino acid residues at positions i and j along the chain is the sum of the interaction energies between the atoms of the respective residues. When e_{ij} is plotted as a function of i, the graph represents the energy distribution of a sequence pair in terms of sequence positioning [18].

3 Results

3.1 Identification of the bHLH motif

3.1.1 Multiple alignments

The most powerful method available today for inferring the biological function of a protein from its sequence is similarity searching on a protein sequence database. Sequence comparison showed that the bHLH motif is present in many proteins. First, we aligned the sequence of each HIF-1α and ARNT subunit along with the partial sequences of other bHLH factors of known 3D structures, using CLUSTALW [5]. Such multiple alignment was designed to detect a portion of the query sequence that would correspond to a similar secondary and tertiary structure, in this case the bHLH motif.

As expected, both subunits possess a bHLH motif. Essential residues are conserved among these sequences. Comparisons of the sequences show four identical hydrophilic residues conserved within the N-terminal end of the "homologous" basic region and a set of hydrophobic residues in two segments, which are separated by a short sequence (Table 1). The basic region is mainly composed of Arg and Lys residues. The two segments could correspond to the two helices found in each monomer. In these segments a clear pattern of conserved and non-conserved residues is found. The conserved residues are separated by three or four residues, which suggests turns of an α-helical protein. In Helix 1, the number of hydrophobic residues is greater but this could account for the dimerization, which is driven by hydrophobic interactions.

3.1.2 Motif detection

Nevertheless, the alignments were not reliable enough because Table 1 shows that some portions of the

Table 1. Partial sequence alignments of the DNA binding domain of ARNT and HIF-1α with Max, USF [25], e47 [26] and MyoD [20]

```
                          Amphipatic Helix 2                    Amphipatic Helix 1
Max------ADKRAHHNALERKRRDHIKDSFHSLRDSVP--SLQGEK--ASRAQILDKATEYIQYMRRKNH
USF------EKRRAQHNEVERRRRDKINNWIVQLSKIIPDCSMESTKSGQSKGGILSKACDYIQELRQSNH
e47------RERRMANNARERVRVRDINEAFRELGRMCQ-MHLKSDK-AQTKLLILQQAVQVILGLEQQ--
Myod-----ADRRKAATMRERRRLSKVNEAFETLKRCTS--SNPNQR--LPKVEILRNAIRYIEGLQALLR
Arnt----ERLARENHSEIERRRRNKMTAYITELSDMVPTCSALARK--PDKLTILRMAVSHMKSLRGTGN
      87                                                              147

Max------ADKRAHHN-ALERKRRDHIKDSFHSLRDSVP--SLQGEK--ASRAQILDKATEYIQYMRRKNH
USF------EKRRAQHN-EVERRRRDKINNWIVQLSKIIPDCSMESTKSGQSKGGILSKACDYIQELRQSNH
Myod-----ADRRKAAT-MRERRRLSKVNEAFETLKRCTS--SNPNQR--LPKVEILRNAIRYIEGLQALLR
e47------RERRMANN-ARERVRVRDINEAFRELGRMCQ-MHLKSDK-AQTKLLILQQAVQVILGLEQQ--
HIF ---ISSERRKEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGD
     13                                                              78
```

sequences were not conserved and therefore the global alignment was probably not optimal. In order to further confirm those results, we used the program GIBBS [8] that detects common motifs between sequences of different proteins. Even though GIBBS and sequence alignments are two different techniques, Table 2 shows a complementarity between the results obtained by both methods. In fact, the Motif 1 detected by GIBBS corresponds to the amphipatic Helix 1 found by alignment (Table 1). Similarly, the sequence portion attributed to Helix 2 (Table 1) is found to be Motif 2 (Table 2). For illustration, we represent the two motifs detected for only five sequences (Table 2).

These two programs allowed us to have a good indication of the residues involved in the bHLH motif and gave two "target" sequences, linked by a loop, which are represented in Table 3 and which will be used in the rest of the work.

In order to validate the target sequences we submitted both sequences to a secondary structure prediction algorithm, PHD [9, 10] (see Methods for description). The results obtained by PHD are shown in Table 4. The secondary structure elements corresponding to the sequence portions are represented by a one letter code (H = helix) and blanks represent secondary structures such as loops. The reliability index of predicition, ranging from 0 to 9, is also represented. The comparison between the target sequence of ARNT in Table 3 and the results obtained by PHD, for that same sequence, show a great complementarity. In fact, PHD detects two helices and a loop region. This region corresponds exactly to the one proposed in Table 3. Concerning the helices, the similarity is evident and we see that the reliability index of predicition is quite high, except for the N-terminal part of the sequence where no secondary structure has been attributed (Table 4). However, the complementarity is not as evident for the target sequence

Table 2. bHLH motifs detected with GIBBS

Motif 2		Motif 1
KRAHHNALERKRRDHIKDSFHSLRDSVP	Max	ILDKATEYIQYMRRKN
RRKAATMRERRRLSKVNEAFETLKRCTS	Myod	ILRNAIRYIEGLQALL
RRAQHNEVERRRRDKINNWIVQLSKIIP	USF	ILSKACDYIQELRQSN
RKEKSRDAARSRRSKESEVFYELAHQLP	**HIF-1α**	**RLTISYLRVRKLLDAG**
ARENHSEIERRRRNKMTAYITELSDMVP	**ARNT**	**ILRMAVSHMKSL**

of HIF-1α. Nevertheless, PHD predicts a helix (Amphipatic Helix 1) with a high reliability between residues VMRL...RVR. This helical structure corresponds to one helix proposed in Table 3. The results do not show a clear loop region for this target sequence, but the reliability index of predicition in this region is low, therefore the presence of a loop region in that sequence portion is not rejected. For the Amphipatic Helix 2 (Table 4), the program predicts part of it as a helix but is not reliable for the N-terminal part of its related sequence. Even though the complementarity is not as evident, one can still see a similarity between the results obtained by PHD (Table 4) and the ones proposed in Table 3.

The secondary structure predictions show that the target sequences obtained by sequence alignments and by the program GIBBS correspond to helix-loop-helix motifs, as previously proposed. These results encourage us to go further in the construction of the model.

3.2 Generation of a 3D model

We then made sequence/structure alignments on the target sequences in order to find a template for the 3D model. The program BLAST [11] was used to align the query sequence among sequences of a database (PDB) from which the 3D structure has been obtained either by X-ray or NMR. In our case, the use of BLAST is justified by the fact that the aim of this part is to find a template but it is important to note that other methods are available such as PSI-BLAST [27] to detect weak similarities. The results obtained by BLAST are represented in Table 5. For the target sequence HIF-1α the top matching sequence corresponds to the bHLH domain of the transcription factor MyoD [20]. The alignment obtained by the program is also shown. Therefore, we will use this structure as a template for the construction of the HIF-1α subunit. For the second target sequence, the first seven hits did not correspond to a bHLH transcription factor but the eigth hit did. In fact, this hit corresponds to the transcription factor MyoD as proposed for the HIF-1α target sequence, so we discarded the first seven results since their structures were not what we expected and were rather different from the bHLH transcription factors. Again, we chose the 3D structure of MyoD as a template to build the ARNT subunit.

Table 3. Target sequences (lower case = α-helix; upper case = loops)

```
HIF-1α : serrkeksrdaarsrrskesevfyelahqLPLPHNVSSHLDKASvmrltisylrvrkll

ARNT :   erlarenhseierrrrnkmtayitelsdmVPTCSALARKPDKLTilrmavshmkslrgt
```

Table 4. Secondary structure prediction obtained by PHD (loops are in bold)

```
                Amphipatic Helix 2                          Amphipatic Helix 1
HIF-1α SERRKEKSRDAARSRRSKESEVFYELAHQ**LPLPHNVSSHLDKAS**VMRLTISYLRVRKLL
          HHHHHH    HHHHHHHHHH    HHHHHHH HHHHHHHHHHHHHHHH
       99987897426467301116999999998689625687534658999999999788449

                Amphipatic Helix 2                          Amphipatic Helix 1
ARNT   ERLARENHSEIERRRRNKMTAYITELSDM**VPTCSALARKPDKLT**ILRMAVSHMKSLRGT
          HHHHHHHHHHHHHHHHHHHH                  HHHHHHHHHHHHH
       98865563467887899999999999996587499799999885359999999998559
```

Table 5. Sequence/structure alignment results obtained by BLAST

HIF-1α

```
TOP 10 matching sequences

>1MDY TRANSCRIPTION ACTIVATION/DNA
>1ACM TRANSFERASE (CARBAMOYL-P,ASPARTATE
>2CPP OXIDOREDUCTASE(OXYGENASE)
>1HLE HYDROLASE INHIBITOR(SERINE PROTEIN
>1PHS PLANT SEED STORAGE PROTEIN (VICILI
>2ASR CHEMOTAXIS
>1TRK TRANSFERASE(KETONE RESIDUES)
>1TRK TRANSFERASE(KETONE RESIDUES)
>1PCN LIPASE PROTEIN COFACTOR
>3RP2 SERINE PROTEINASE

sequence alignement produced by BLAST

Hit 1 >1MDY TRANSCRIPTION FACTOR (MyoD)
HIF-1α:   1 SERRKEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYL 53
            ...:.....:.:.:.:.:.:.....:....:.:....:..:.:.
MyoD:     9 NADRRKAATMRERRRLSKVNEAF-ETLKRS-TSSNPNQRLPKVEILRNAIRYI 59
```

ARNT

```
TOP 10 matching sequences

>3PGM TRANSFERASE (PHOSPHORYL)
>1RBL LYASE(CARBON-CARBON)
>1ALA CALCIUM/PHOSPHOLIPID-BINDING PROTE
>1ULA PENTOSYLTRANSFERASE
>1IDS SUPEROXIDE DISMUTASE
>1CGT GLYCOSYLTRANSFERASE
>1PYA CARBOXY-LYASE
>1MDY TRANSCRIPTION ACTIVATION/DNA
>1COM CHORISMATE MUTASE
>1BST HORMONE

sequence alignement produced by BLAST

Hit 8 >1MDY TRANSCRIPTION FACTOR (MyoD)
ARNT:     1 ERLARENHSEIERRRRNKMTAYITELSDMVPTCSALARKPDKLTILRMAVSHMKSLR 57
            ....:.....:::.....:.....:.:.....:..:::.:.....:.
MyoD:     8 TNADRRKAATMRERRRLSKVNEAFETLKR-STSSNPNQRLPKVEILRNAIRYIEGLQ 63
```

We then substituted the residues of HIF-1α and ARNT, corresponding to two helical structures, into the 3D coordinates of the template's residues. Once the two helices for each monomer were constructed, we generated the loops by means of HOMOLOGY [13]. We could not substitute the sequence portion corresponding to the loops onto the template for two reasons. Firstly, the sequence alignments did not show any conservation or any evident patterns in the loop portion. Secondly, loops are known to be more flexible and less structured regions. The conformations generated for the sidechains of the residues in generated loops are arbitrarily set to be fully extended. This is not a real situation and is never observed in proteins. Therefore, each loop was submitted to the SCWRL [12] program. Each residue of the sidechains was placed in an optimized conformation, based on the backbone-dependent library. Figure 1 represents the structures of ARNT and HIF-1α obtained in this way.

Thirdly, we minimized the structure of each subunit separately, using DISCOVER [16]. These energy optimizations were mainly used to avoid steric clashes. However, after several steps of energy minimization, the structure of each subunit started to diverge from a bHLH structure. Instead of staying in the HLH structure, as in Fig. 1, the structure of each monomer became distorted. Each helix was broken into little fragments. Moreover, the Helix 1 had a tendancy to tip and face the other helix. This led us to think that the secondary structure (helical in this case) was induced upon dimerization and stabilized by interactions with its DNA consensus sequence.

Therefore, we first assembled both subunits to construct the heterodimer before optimisation (Fig. 2). This was possible since the template we used to build the monomers was initially a dimer. The structure of the heterodimer was then optimized. The results show that the helices responsible for the dimerization now remained well structured (Fig. 2). Nevertheless, the helices responsible for the protein/DNA interaction were not as well structured. Once again, this could be explained by the fact that the DNA sequence recognized by HIF-1

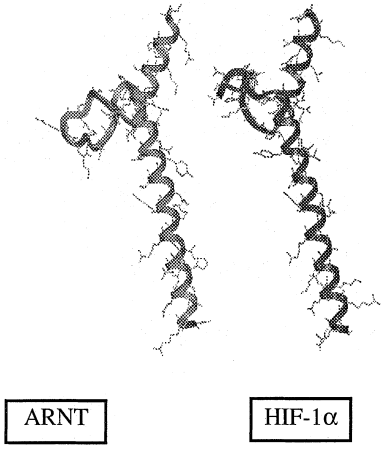

Fig. 1. Representation of the 3D structure of the bHLH domain of ARNT and HIF-1α

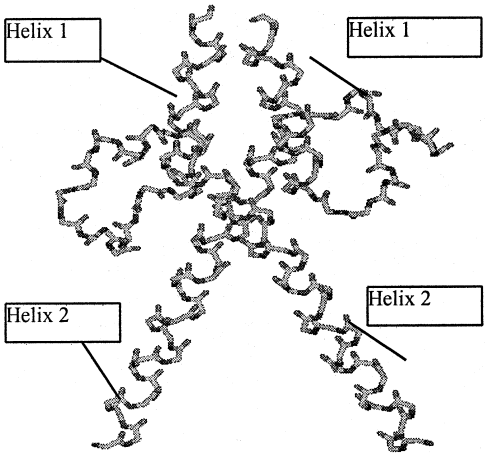

Fig. 2. Representation of the 3D structure of the heterodimer (bHLH domains of HIF-1)

was not taken into account in this calculation. Recently, Fisher et al. [21] have tested, by circular dichroism, whether the basic region of the b-HLH-ZIP proteins might similarly form a helix upon binding. Their results confirmed that the helical structure was induced when the protein binds DNA.

3.3 Analysis of the structure

We used different tools such as Ramachandran plots, energetic profiles, and physicochemical properties to test if the model obtained was plausible. The energy minimization procedures gave four different models, including the initial one. Each of these models corresponded to a different "derivative criterion (DC)" used in the energy minimization steps.

We calculated the Ramachandran plot for three models. The first one is the initial heterodimer 3D model (no minimization), the second one is an intermediate model (DC = 0.1 kcal/Å), and the last model the one obtained once the minimization procedure was finished (DC = 0.01 kcal/Å). The Ramachandran plots for these structures show that the number of residues in disallowed regions decreases to a certain point as the energy minimization goes on. However, for the last structure, the number of residues in disallowed regions was greater. This means that even though the energy of the structure was decreasing, some residues were forced to adopt unfavorable conformations. This result confirms what we observed during the energy minimization procedures: one of the helices was not as well structured probably owing to the absence of the DNA.

We propose as a plausible model of the bHLH the 3D model with the DC of 0.1 kcal/Å, because of its low potential energy and because it has the greatest number of residues in the favored regions. It must be noted that the few residues which are present in disallowed regions are located in the loop regions. This also confirms the difficulty to model this loop region.

We also calculated, on the selected model, the polar and hydrophobic contributions for all the residues, using

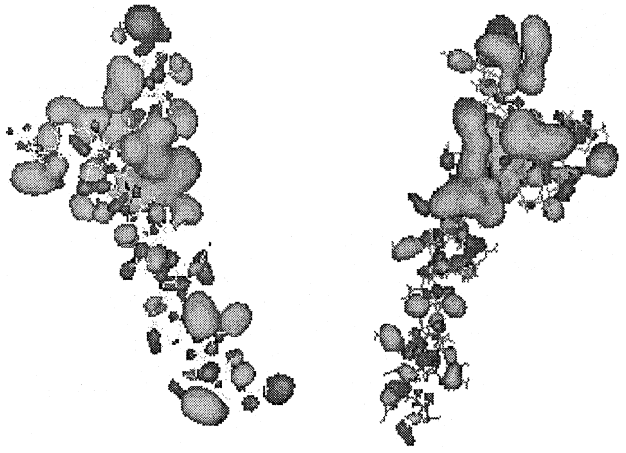

Fig. 3. Representation of polar (*dark gray*) and hydrophobic (*light gray*) contributions of the residues

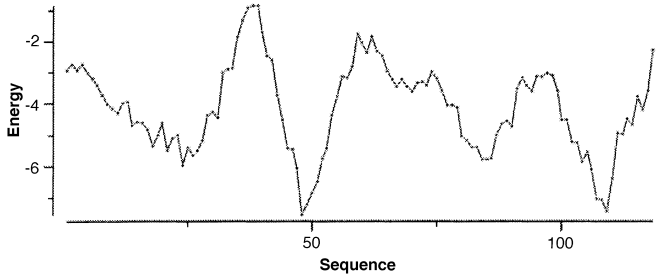

Fig. 4. Energetic profile of the model for the bHLH domain of HIF-1α (units are represented in E/kT)

the program HINT [17]. Figure 3 represents the polar contributions (dark gray) at the −20 level and the hydrophobic contributions (light gray) at the 5 level. The complementarity between the two hydrophobic patches is evident. This complementarity is responsible for the dimerization of both subunits, which is mediated by hydrophobic interactions. This figure shows that the hydrophobic amino acids from Helices 1 and 2 are buried in the interior of the four-helix bundle, where they pack together to assure the dimerization and stability of the heterodimer structure. We also expect a greater polar contribution in the two helices that interact with DNA. In this region, the basic residues make up most of the DNA/protein interactions. Ma et al. [22] reported different types of interactions in that region such as hydrogen bonds, water-mediated contacts between a carboxylate group and a nitrogen atom of a base of the DNA, and salt bridge interactions. Figure 3 clearly shows that both monomers can be brought together through the complementarity of the hydrophobic residues and that the basic region would be inserted into the major groove of the DNA.

To further validate the model, we calculated, using PROSA [19], the energetic profile of the structure (Fig. 4). The energy calculated for each residue is negative, thus most residues are in a favorable conformation. We noticed, however, an increase in energy for the residues in the loop regions. This result is in agreement with the Ramachandran plot.

4 Conclusion

The combination of different modeling approaches leads to the definition of a valuable 3D structure of the bHLH domain of the transcription factor HIF-1. The next logical step in this work would be to dock this structure onto the DNA sequence recognized by HIF-1. We are currently investigating this approach using the program FTDOCK [23]. The docking of the protein/DNA would permit us to characterize the protein/DNA interactions and identify the residues and the bases of the DNA involved in the protein/DNA recognition and interaction. Such a model will have to be confirmed by X-ray analysis of the HIF-1 dimer bound to DNA. It will help to understand the DNA/protein interactions and the role played by HIF in the cellular response to hypoxia.

Acknowledgements. C.M. is Research Associate of the FNRS (National Funds for Scientific Research, Brussels, Belgium). E.M. is grateful to the FRIA for financial support.

References

1. Wang GL, Jiang B, Semenza GL (1995) Biochem Biophys Res Commun 216:669
2. Kvietikova I, Wenger R, Marti H, Gassmann M (1995) Nucleic Acids Res 23:4542
3. Semenza GL, Wang GL (1992) Mol Cell Biol 12:5447
4. Yuxiang L, Shanna RC, Toshisuke M, Stelle K (1995) Circ Res 77:638
5. Semenza GL, Roth PH, Fang H, Wang GL (1994) J Biol Chem 269:23757
6. Wang GL, Jiang B, Rue EA, Semenza GL (1995) Proc Natl Acad Sci USA 92:5510
7. Higgins DG, Sharp PM (1998) Gene 73:237
8. Lawrence CE, Altshul SF, Boguski MS, Liu JS, Neuwald AF, Wooton JC (1993) Science 268:208
9. Rost B, Sander C (1993) J Mol Biol 232:584
10. Rost B, Sander C (1994) Proteins 19:55
11. Altschul SF, Gish W, Miller W, Myers E, Lipman DJ (1991) J Mol Biol 215:403
12. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr., Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) J Mol Biol 112:535
13. Homology User Guide (1996) Molecular Simulation, San Diego, Calif.
14. Bower MJ, Cohen FE, Dunbrack RL Jr (1997) J Mol Biol 267:1268
15. Discover User Guide (1996) Molecular Simulation, San Diego, Calif.
16. Kelllog GE, Serrus SF, Abraham DJ (1991) Comput Aided Mol Design 5:545
17. Hansch C, Leo AI (1979) Substituent constants for correlation analysis in chemistry and biology. Wiley, New York
18. Sippl MJ (1993) Proteins Struct Funct Genet 17:355
19. Sippl MJ (1990) J Mol Biol 213:859
20. Braun T, Bober E, Buschhausen-Denker G, Kotz S, Grzeschik K-H, Arnold HH (1989) EMBO J 8:3617
21. Fisher DE, Parent LA, Sharp PA (1993) Cell 12:467
22. Ma PC, Rould MA, Weintraub H, Pabo CO (1994) Cell 77:451
23. Jackson RM, Gabb HA, Sternberg MJ (1998) J Mol Biol 276:265
24. Blackwood EM, Eisenman RN (1991) Science 251:1211
25. Gregor PD, Sawadogo M, Roeder RG (1990) Genes Dev 4:1730
26. Murre C, McCaw PS, Baltimore D (1989) Cell 56:777
27. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Nucleic Acids Res 25:3389