Molecular modeling of the three-dimensional structure of the human interleukin-11

Ei-Rong Chenb, Ting-Jun Houa, Xiao-Jie Xua*, Zheng Lib and Jia-Quan Wanga

aChemistry Department, Beida-Jiuyuan Molecular Design Laboratory, Peking University, Beijing, 100871; bDepartment of Technique-Physics, Peking University, Beijing, 100871, China

Introduction

Interleukin-11 (IL-11) is a multifunctional cytokine possessing wide biological activities which can affect growth differentiation in several hematopoietic cell types, including early pluripotent stem cells, megakaryocyte progenitors and megakaryocytes, erythrocyte progenitors, etc. [1]. It is also identified as inhibiting adipogenesis in preadipocytes and stimulating production of several acute phase plasma proteins in hepatocytes. IL-11 has been shown to share many biological activities with IL-6, including the ability to activate expression of the same early response genes.

Until now, the 3D-structure of IL-11 has not been available from experiments, although large studies have been performed. From previous studies, it has been revealed that the cytokine family shares a common structural motif consisting of four anti-parallel helices in an up-up-down-down configuration. To date, the four-helix bundle fold has been observed in human growth hormone (HGH), interleukin-2 (IL-2), interleukin-4 (IL-4), granulocyte-colony-stimulating factor (G-CSF), inter-leukin-5 (IL-5) and human macrophage colony-stimulating factor (GM-CSF). The cytokine family is one of the most diverse protein families but their structural relationships can be determined by comparing their amino acid sequences. From sequence alignment, it can be found that only G-CSF shares sequence similarity higher than 30% with IL-11.

In this paper, we have combined sequence comparison and secondary structural prediction to generate a molecular model of interleukin-11. The X-ray structure of G-CSF [2] was used as a template to produce the model, moreover, in the sequence alignment, the information from the secondary structural prediction of the GOR_II method [3] was taken into account. The validation of the model suggests that the quality of this structure is sound. The predicted model may aid in the design of experiments to better understand IL-11 and its interaction with receptors.

Results and Discussion

A pair-wise sequence alignment of IL-11 and G-CSF was produced using the Needleman and Wunsch algorithm [4]. When performing the sequence alignment, the penalty of a gap was defined to be proportional to the length of the gap. GOR_II
method was applied to predict the secondary structure of IL-11. The sequence alignment of IL-11 and G-CSF was modified manually to adjust the structurally significant region of IL-11 according to information from the secondary structure prediction. Gaps were avoided in the predicted continuous helical regions. For model building of IL-11, the coordinates of the sequences of conserved regions came directly from the template protein. The coordinates for loops were assigned by searching the Brookhaven Protein Databank (PDB). In these calculations, a standard rotamer library was used and a cutoff distance of 10Å was applied for treatment of nonbonded interaction. In addition, the initial building model was energy minimized with CVFF force field until the root square mean (rms) derivative of the energy was less than 0.1 kcal/ (mol. Å). The completed model was validated using the Health program in Quanta and was checked from several aspects. Energy minimization calculations were carried out using the DISCOVER program in INSIGHTII (MSI, San Diego, CA, U.S.A), and the prediction of the IL-11 model was performed using the Homology program in INSIGHTII.

From the sequence database search we found that a total of three proteins shared a similarity larger than 30% with IL-11. But from the folding check of their 3-D structures, only G-CSF belongs to the same family as IL-11. G-CSF not only shares relative high sequence similarity with IL-11, but also possesses the same structural motif. Thus the X-ray structure of G-CSF was used as a reference protein to model the 3D-structure of IL-11. GOR-II method was carried out to predict the secondary structure of IL-11, but the contributions of the helical region had to be considered. The alignment was carefully modified manually so that the gaps appearing in the continuous helix were as few as possible. Fig. 1 shows the sequence alignment of G-CSF and IL-11 after manual adjustment.

Based on the alignment, the three-dimension model was constructed using the X-ray structure of G-CSF as the template. The coordinates of the sequences of the reserved regions were obtained from G-CSF except for those residues from 164-178. The analysis from the sequence alignment and secondary structure prediction suggested residues 149-174 should be a continuous helix, but sequence alignment introduced a gap between them that separates the helix into two parts. It is thus not appropriate to copy coordinates from G-CSF for this segment. For these residues from 164 to 178, the 3-D structures instead are modeled according to the standard structure of an α-helix. Intermolecular contacts and stereochemistry of the initial assembled model was refined by energy minimization using DISCOVER program.