Published in final edited form as:

Int J Quantum Chem. 2008 ; 108(5): 1017–1021. doi:10.1002/qua.21553.

Evaluation of methods to cap molecular fragments in calculating energies of interaction in avian pancreatic polypeptide

Marcus P. D. Hatfield^{a,b}, Nicholas Y. Palermo^b, József Csontos^b, Richard F. Murphy^b, and Sándor Lovas^{b,1}

^a Nebraska Wesleyan University, Lincoln, NE 68504, U.S.A

^b Department of Biomedical Sciences, Creighton University, Omaha, NE 68178, U.S.A

Abstract

The accuracy of the determination of the energy of interaction between Phe20 and the Pro5-Thr6-Tyr7-Pro8 complex inside the hydrophobic core of avian pancreatic polypeptide was investigated using three capping strategies for molecular fractionation with conjugated caps and DFT quantum chemical calculations at the BHandHLYP/cc-pVTZ level of theory. The most accurate determination resulted from acetylation of the α -amino group combined with methyl amidation of the α -carbonyl group with relative deviations less than 10%. Combinations of hydrogenation of the α -amino group with the replacement of the α -carbonyl group with a hydrogen and the hydrogenation of the α -amino group with methylation of the α -carbonyl group were less accurate, leading to relative deviations up to 35%. Choice of capping methods depends on the structural features of the polypeptide system, the desired accuracy and the available computational resources.

Keywords

density functional theory; weakly polar interactions; interaction energies; avian pancreatic polypeptide; molecular fractionation with conjugated caps

Introduction

Non-covalent interactions play important role in stabilization tertiary fold of proteins. Molecular dynamics simulations are routine tools for characterization of geometries of interacting residues, however, quantum chemical calculations are needed to determine exact energies of such interactions. Molecular fractionation with conjugate caps (MFCC) has been used for the calculation of energies of interaction of protein with water,^{1–3} drugs^{4, 5} and ligands^{6, 7} and DNA with ligands.⁸ In this approach a polypeptide (Figure 1) is conceptually fragmented into partial structures which are then terminally capped to preserve the original local electronic environment that was removed.^{1, 9} Capping groups should be as small as possible to minimize computation time.¹⁰

In this study the accuracy of using: **1**, acetylation of the α -amino group with methyl amidation of the α -carbonyl group (Figure 2 A–D); **2**, hydrogenation of the α -amino group with the replacement of the α -carbonyl group with a hydrogen (Figure 3 A–D) and **3**, hydrogenation of the α -amino group with the methylation of the α -carbonyl group (Figure 4 A–D) as capping

¹Corresponding author: Sándor Lovas, Department of Biomedical Sciences, Creighton University, 2500 California Plaza, Omaha, NE 68178, Phone: 402-280-5753, Fax: 402-280-2690, E-mail: slovas@creighton.edu.

strategies to calculate energies of interaction of structures in the hydrophobic core of avian pancreatic polypeptide (aPP) was compared.

aPP¹¹ (Figure 5), Gly-Pro-Ser-Gln-Pro⁵-Thr-Tyr-Pro-Gly-Asp¹⁰-Asp-Ala-Pro-Val-Glu¹⁵-Asp-Leu-Ile-Arg-Phe²⁰-Tyr-Asp-Asn-Leu-Gln²⁵-Gln-Tyr-Leu-Asn-Val³⁰-Val-Thr-Arg-His-Arg³⁵-Tyr-NH₂, is a 36 residue polypeptide which is a member of the pancreatic polypeptide hormones,¹² peptide YY¹³ and neuropeptide Y¹⁴ family peptides. It regulates gastrointestinal and cardiovascular function, and increases food intake.¹⁵ The secondary structure of aPP consists of a poly-proline II helix (residues 1–8) a β -turn (residues 9–13), and an α -helix (residues 14–32).¹⁶ The poly-proline II helix and α -helix run anti-parallel and their non-polar groups pack together to form a hydrophobic core,^{16, 17} which is stabilized primarily by the weakly polar interactions.¹⁸ Weakly polar interactions in the hydrophobic core of aPP include the CH- π interactions of Tyr7 with Pro8 and Phe20, Phe20 with Pro5, Pro8 and Leu24, Tyr21 with Gln25 and Tyr27 with Pro2, Gln4, Pro5 and Val31, an Ar-NH interaction between Tyr7 and Gly9 and an Ar-Ar interaction between Tyr7 and Phe20. It was, thus, possible to show that the total stabilization energy of these interactions, recently determined here, is $-27.33 \text{ kcal mol}^{-1}$.¹⁹

Methods

Molecular Fragmentation

The structure of aPP (2BF9) was obtained from the Protein Data Bank²⁰. The polypeptide was fragmented at the peptide bonds between Gln4 and Pro5, Pro5 and Thr6, Thr6 and Tyr7, Tyr7 and Pro8, Pro8 and Gly9, Arg19 and Phe20 and Phe20 and Tyr21. The accuracy of three different capping techniques (**1**, **2** and **3**; see above) was tested as follows. The energy of interaction of Phe20 with the Pro5-Thr6-Tyr7-Pro8 fragment was calculated directly (ΔE_{int}) and with individual Pro5, Thr6, Tyr7 and Pro8 residues by the MFCC method ($\Delta E_{\text{int}}[\text{MFCC}]$) and the deviation (D_{abs}) and relative deviation (D_{rel}) were determined:

$$\begin{aligned} \Delta E_{\text{int}}[\text{MFCC}] &= \Delta E_{\text{int}}[\text{Phe20} - \text{Pro5}] + \Delta E_{\text{int}}[\text{Phe20} - \text{Thr6}] + \\ &\Delta E_{\text{int}}[\text{Phe20} - \text{Tyr7}] + \Delta E_{\text{int}}[\text{Phe20} - \text{Pro8}] - \\ &\Delta E_{\text{int}}[\text{Phe20} - \text{CC}_{\text{Pro5Thr6}}] - \Delta E_{\text{int}}[\text{Phe20} - \text{CC}_{\text{Thr6Tyr7}}] - \\ &\Delta E_{\text{int}}[\text{Phe20} - \text{CC}_{\text{Tyr7Pro8}}] \end{aligned} \quad (1)$$

where CC_{XY} is the conjugated caps of residues X and Y (Figures 2D, 3D and 4D);

$$D_{\text{abs}} = \Delta E_{\text{int}} - \Delta E_{\text{int}}[\text{MFCC}] \quad (2)$$

and

$$D_{\text{rel}} = D_{\text{abs}} / \Delta E_{\text{int}}. \quad (3)$$

For pairs of interacting fragments, combinations of capping strategies were **1, 1**; **1, 2**; **1, 3**; **2, 1**; **2, 2**; **2, 3**; **3, 1**; **3, 2** and **3, 3**. The first numbers represent the capping strategy used for Phe20. The second numbers represent capping strategy used for either Pro5-Thr6-Tyr7-Pro8 fragment or the individual Pro5, Thr6, Tyr7 and Pro8 residues. The positions of all atoms added during capping were optimized using the B3LYP/6-31G** level of theory.

Interaction energy calculations

Energies of interaction of the molecular fragments were calculated at the BHandHLYP/cc-pVTZ level of theory. This functional was chosen because it closely approximates high-level results for weakly polar interactions and dispersion systems.^{19, 21–24} The BHandHLYP functional is expressed as follows:

$$0.5 \cdot E_x^{HF} + 0.5 \cdot E_x^S + 1.0 \cdot E_c^{LYP}$$

where E_x^{HF} is the exact Hartree-Fock exchange, E_x^S is the exact Slater exchange, and E_c^{LYP} is the Lee-Yang-Parr correlation. The counterpoise procedure was used to correct the basis set superposition error (BSSE).²⁵

Programs

The Jaguar V5.5 release 11 and V6.0 release 11 (Schrödinger LLC, Portland, OR) program package was used for all calculations. YASARA (<http://www.yasara.org>) was used for visualization and the preparation of figures.

Results and Discussion

The MFCC approach led to an overestimate of the energy of interaction in seven of the nine combinations of capping strategies (Table I). Nevertheless, the absolute deviations were less than 1 kcal mol⁻¹ except for the **2, 2**; **2, 3** and **3, 3** combinations. When capping strategy **1** is applied to residues 5–8 the relative deviation of the energy of interaction is close to -9.6% except when capping strategy **2** is used for Phe 20 when it is substantially lower at -7.92% (Table II). When capping strategy **2** is used for residues 5–8 the relative deviation of the energy of interaction varies between 8.17 and -24.66%. This is similar to the results for capping strategy **3** which has relative deviations in the energies of interaction between 8.84 and -34.38%. The greatest accuracy is obtained with combination **2, 1** with the model structures here. The relative deviations for combinations **1, 1**; **2, 2** and **3, 3**, respectively, are -9.59, -24.66 and -22.95% and a computational cost of 338, 654 and 212 hours. Capping strategy **3** is used frequently^{26, 27} for polypeptides in general but should be avoided because it can lead to less accurate determination of the energy of interaction. Capping strategy **2** leads to similar accuracy to that achieved by **3** but is 3.1 times more computationally expensive. Thus capping strategy **1** appears to be the most appropriate for calculating the energies of interaction between residues in the hydrophobic core of the chosen model polypeptide because it facilitated the most accurate determination of the energy of interaction and has a computational cost of only 1.6 times that of capping strategy **3**. The findings here are based on the interactions inside a hydrophobic core where the distance between fragments is small. This often gives rise to larger errors using the MFCC method.¹ Thus, these results may not be applicable to systems in which interacting fragments are either further apart or in an aqueous environment. When choosing a capping strategy it is important to consider the individual system under study and the desired accuracy. Strategies which introduced interactions between newly introduced hydrogens and native structures should be avoided. In general the capping strategies should be comparable with the recent adaptations of MFCC including EC-MFCC,²⁸ EFA-MFCC²⁹, MFCC-DM³⁰, MFCC-DM-PIC³¹ and MFCCII.¹⁰

Conclusions

Capping strategy **1** facilitated the most accurate determination of energy of interaction of Phe20 with the Pro5, Thr6, Tyr7 and Pro8 complex using the MFCC method, with D_{rel} less than 10%. Capping strategies **2** and **3** lead to much larger D_{rel} . Optimal capping strategies should best describe the local electronic environment while introducing the fewest atoms. Choice of capping methods depends on the structural features of the polypeptide system, the desired accuracy and the available computational resources. Therefore, capping strategies should be tested for each system.

Acknowledgements

This work was supported by the NIH-INBRE grant P20 RR016469 and the Carpenter Endowed Chair in Biochemistry, Creighton University.

References

1. Zhang DW, Chen XH, Zhang JZH. *J Comput Chem* 2003;24:1846–1852. [PubMed: 14515367]
2. Mei Y, Ji C, Zhang JZ. *J Chem Phys* 2006;125:094906–094913. [PubMed: 16965118]
3. Chen XH, Zhang DW, Zhang JZ. *J Chem Phys* 2004;120:839–844. [PubMed: 15267920]
4. He X, Mei Y, Xiang Y, Zhang DW, Zhang JZ. *Proteins* 2005;61:423–432. [PubMed: 16114038]
5. Mei Y, He X, Xiang Y, Zhang DW, Zhang JZ. *Proteins* 2005;59:489–495. [PubMed: 15789428]
6. Zhang DW, Xiang Y, Gao AM, Zhang JZ. *J Chem Phys* 2004;120:1145–1148. [PubMed: 15268233]
7. Xiang Y, Zhang DW, Zhang JZ. *J Comput Chem* 2004;25:1431–1437. [PubMed: 15224387]
8. Chen XH, Zhang JZ. *J Chem Phys* 2004;120:11386–11391. [PubMed: 15268172]
9. Zhang D, Zhang JZH. *J Chem Phys* 2003;119:3599–3605.
10. Mei Y, Zhang DW, Zhang JZH. *J Phys Chem A* 2005;109:2–5. [PubMed: 16839082]
11. Kimmel JR, Hayden J, Pollock HG. *J Biol Chem* 1975;250:9369–9376. [PubMed: 1194289]
12. Inui A, Miura M, Teranishi A, Hirose Y, Nakajima M, Okita M, Nishimura R, Baba S, Kasuga M. *Endocrinology* 1992;130:3314–3322. [PubMed: 1597144]
13. Tatemoto K, Mutt V. *Nature* 1980;285:417–418. [PubMed: 6892950]
14. Tatemoto K, Carlquist M, Mutt V. *Nature* 1982;296:659–660. [PubMed: 6896083]
15. Denbow DM, Duke GE, Chaplin SB. *Peptides* 1988;9:449–454. [PubMed: 3420004]
16. Blundell TL, Pitts JE, Tickle IG, Wood SP, Wu CW. *Biochemistry* 1981;20:4175–4179.
17. Glover ID, Moss DS, Tickle IJ, Pitts JE, Haneef I, Wood SP, Blundell TL. *Adv Biophys* 1985;20:1–12.
18. Zhang H, Wong CF, Thacher T, Rabitz H. *Proteins: Structure, Function, and Genetics* 1995;23:218–232.
19. Palermo NY, Csontos J, Murphy RF, Lovas S. *Int J Quantum Chem* 2008;108:814–819.
20. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. *Nucleic Acids Research* 2000;28:235–242. [PubMed: 10592235]
21. Palermo NY, Csontos J, Owen MC, Murphy RF, Lovas S. *J Comput Chem* 2007;28:1208–1214. [PubMed: 17299770]
22. Waller MP, Robertazzi A, Platts JA, Hibbs DE, Williams PA. *J Comput Chem* 2006;27:491–504. [PubMed: 16444702]
23. Piacenza M, Grimme S. *J Comput Chem* 2004;25:83–99. [PubMed: 14634996]
24. Romero C, Fomina L, Fomine S. *Int J Quantum Chem* 2005;102:200–208.
25. Boys SF, Bernardi F. *Mol Phys* 1970;19:553–566.
26. Klusak V, Havlas Z, Rulisek L, Vondrasek J, Svatos A. *Chemistry & Biology* 2003;10:331–340. [PubMed: 12725861]
27. Vondrasek J, Bendova L, Klusak V, Hobza P. *J Am Chem Soc* 2005;127:2615–2619. [PubMed: 15725017]
28. Li S, Li W, Fang T. *J Am Chem Soc* 2005;127:7215–7226. [PubMed: 15884963]
29. Jiang N, Ma J, Jiang Y. *J Chem Phys* 2006;124:114112–114121. [PubMed: 16555879]
30. Chen X, Zhang Y, Zhang JZ. *J Chem Phys* 2005;122:184105–184110. [PubMed: 15918692]
31. Chen XH, Zhang JZ. *J Chem Phys* 2006;125:044903–044910.

Figure 1.
Polypeptide chain.

Figure 2.

Capped fragments and conjugate caps in capping strategy **1**. N- and C-terminal caps are red and green, respectively. **A**, *N*-Ac-j-k-NHCH₃ fragment; **B**, *N*-Ac-j-NHCH₃ fragment; **C**, *N*-Ac-k-NHCH₃ fragment; **D**, *N*-Ac-NHCH₃ conjugate caps.

Figure 3.

Capped fragments and conjugate caps in capping strategy **2**. N- and C-terminal caps are red and green, respectively. **A**, H-j-k-NHCR'''H₂ fragment; **B**, H-j-NHCR/H₂ fragment; **C**, H-k-NHCR'''H₂ fragment; **D**, H-NHCR'''H₂ conjugate caps.

Figure 4.

Capped fragments and conjugate caps in capping strategy **3**. N- and C-terminal caps are red and green, respectively. **A**, H-j-k-CH₃ fragment; **B**, H-j-CH₃ fragment; **C**, H-k-CH₃ fragment; **D**, H-CH₃ conjugate caps.

Figure 5. Structure of the aPP backbone and interacting atoms of residues Pro5, Thr6, Tyr7, Pro8 and Phe20. The α -helix is blue, β -turn is green and the poly-proline II helix is cyan.

Table I

Interaction of variously capped fragments.

Capping strategy [*]	$\Delta E_{\text{int}}/\text{kcal mol}^{-1}$	$\Delta E_{\text{int}}[\text{MFCC}]/\text{kcal mol}^{-1}$	$D_{\text{abs}}/\text{kcal mol}^{-1}$	CPU Time/h
1, 1	-8.45	-9.26	0.81	338
1, 2	-7.49	-8.35	0.86	357
1, 3	-7.24	-6.60	-0.64	276
2, 1	-8.46	-9.13	0.67	593
2, 2	-6.69	-8.34	1.65	654
2, 3	-7.30	-9.81	2.51	462
3, 1	-7.29	-7.99	0.70	266
3, 2	-7.59	-6.97	-0.62	292
3, 3	-7.32	-9.00	1.68	212

^{*}
See methods.

Table IIPercent D_{rel} for variously capped interacting structures.

Residues 5–8				
Phe20	1	1	2	3
	2	–9.59	–11.48	8.84
	3	–7.92	–24.66	–34.38
		–9.60	8.17	–22.95