

TÍNH TOÁN DỰ BÁO ẢNH HƯỞNG CỦA ĐỘT BIẾN GEN ĐẾN HOẠT TÍNH CỦA CD38 ĐỐI VỚI PHẢN ỨNG TẠO THÀNH CADPR

Đến Tòa soạn 17-8-2007

ĐẶNG ỨNG VẬN

Trung tâm Ứng dụng tin học trong hóa học, Đại học Quốc gia Hà Nội

SUMMARY

A suitable calculation procedure has been introduced to predict the mutagenesis effect on NAD cyclizing activities of CD38, a multifunctional ectoenzyme responsible for catalyzing the conversions from nicotinamide adenine dinucleotide (NAD) to cyclic ADP-ribose (cADPR) and from cADPR to ADP-ribose (ADPR). The results indicate that there are obviously differences between binding energy of NMN on CD38 and on its mutants. That is why all the mutants exhibited enhanced cyclase activity. There is also a linear relationship between binding energy of nicotine amide on E146-mutants and kinetic constant of cyclizing.

I - MỞ ĐẦU

Trong những công trình trước đây [1, 2], chúng tôi đã công bố những kết quả ban đầu ứng dụng tin học trong nghiên cứu các quá trình gắn kết (docking) phối tử lên protein và DNA. Xu hướng ứng dụng tin học trong nghiên cứu các quá trình hóa sinh đang ngày càng phát triển rộng rãi. Với sự xuất hiện các phần mềm kết hợp cơ học lượng tử (QM) và cơ học phân tử (MM) người ta đã không chỉ dự báo khá chính xác vị trí gắn kết của phối tử mà còn tính toán được giá trị các hằng số thụ động (inhibition constant) chẳng hạn như IC50, trong đó $IC50 = 10^{E_{bind}/5.85}$, E_{bind} là năng lượng tự do gắn kết giữa phối tử và protein:

$$E_{bind} = E_{es} + E_{vdw} - TdS + E_{tor} \quad (1)$$

E_{es} là năng lượng tương tác tĩnh điện, E_{vdw} là năng lượng tương tác Van der Waals, E_{tor} là năng lượng biến dạng phối tử và TdS là phần góp của yếu tố entropy. Khả năng đánh giá tương đối chính xác năng lượng gắn kết đã cho phép chúng ta có thể thực hiện các tính toán dự báo cho các hiện tượng hóa sinh phức tạp. Bài báo này trình bày những kết quả đầu tiên chúng

tôi thu được khi sử dụng các phương pháp hóa tin để tính toán tác động đột biến gen đến hoạt tính xúc tác của protein CD38.

II - ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP TÍNH

CD38 được biết như là một kháng thể can dự vào vật chủ có các chức năng lympho bào bao gồm sự biệt hóa (differentiation), sự tăng sinh (proliferation) và chết tế bào theo chương trình (apoptosis) [3, 4]. Sự hiện diện của nó còn phổ biến trong các mô không tạo huyết (nonhematopoietic). Ngoài chức năng kháng thể, CD38 cũng có nhiều hoạt tính xúc tác [5, 6], không chỉ cho sự thủy phân của NAD (nicotinamid adenin dinucleotid) và các cADPR (ribose ADP- vòng) mà còn vòng hóa NAD và các đồng đẳng của nó như là NGD (nicotinamid guanin dinucleotid) để cho ra cADPR và cGDPR (ribose GDP vòng). Thêm nữa, nó có thể dùng NADP (nicotinamid adenin dinucleotid photphat) và axit nicotinic như là các chất nền (substrate) và xúc tác phản ứng trao đổi bazơ để cho ra NAADP (muối photphat adenin nucleotid của axit nicotinic). Cả hai cADPR và NAADP đều là các thể truyền tín hiệu Ca^{2+} hóa giải

(mediating) sự chuyển động của Ca^{2+} nội bào dự trữ trong một diện rộng các tế bào từ động vật đơn bào, thực vật đến con người [7]. Đã có nhiều thực nghiệm được tiến hành theo hướng sử dụng kỹ thuật đột biến gen trực vị (site-directed mutagenesis) tại một vài tâm hoạt động (active site) của CD38 bao gồm Glu226, Trp125, Trp189, Glu146 và Asp155 để thu được những thay đổi rất mạnh mẽ hoạt tính của CD38 trong các phản ứng thủy phân và vòng hóa NAD[8], thủy phân và tổng hợp NAADP[9] và cADPR[10]. Các thể đột biến được khảo sát trong bài báo này gồm: E146L, E146F, E146D, E146G, E146A, E146K và E146Q bằng cách thay thế axit amin GLU146 lần lượt bằng các axit amin: Leucine (L), Phenylalanine (F), axit aspartic (D), Glycine (G), Alanine (A), Lysine (K) và Glutamine (Q). Các thể đột biến E226G của CD38 (mã số trong ngân hàng dữ liệu protein PDB là 2HCT) với sự thay thế Glu226 bằng Gly, phối tử là NMN (nicotinamide mononucleotide - $C_{11}H_{16}N_2O_8P$ - một mảnh (substrate) của CD38), phức của thể đột biến E226Q của CD38 với cADPR (cyclo-ADPriboz) (mã 2O3Q), phức E226D-cADPR (mã 2O3R), phức E226G-cADPR (mã 2O3S) được sử dụng để kiểm nghiệm phương pháp tính do đã có đầy đủ dữ liệu cấu trúc tinh thể trong PDB.

Quá trình tính toán đòi hỏi 3 thuật toán cốt lõi: i) tính năng lượng gắn kết, ii) thay thế từng axit amin trong tâm hoạt động của protein tự nhiên bằng các axit amin khác (tương đương với kỹ thuật đột biến gen trực vị) và tối ưu hình học các thể đột biến (mutant), iii) gắn kết phối tử lên các thể đột biến để tính toán năng lượng gắn kết của phức được tạo thành. Nhằm nâng cao độ tin cậy của các kết quả tính chúng tôi đã sử dụng phối hợp các phương pháp hiện có, đồng thời thực hiện phép so sánh để lựa chọn được phần mềm tối ưu dùng cho các tính toán. Hiện nay đã có các phần mềm được minh chứng rằng có độ phù hợp cao với thực nghiệm như QUANTUM[11] hoặc GLIDE [12] với những áp dụng kỹ thuật QM/MM gần đây nhất [13]. Phiên bản 3.3 của QUANTUM là phần mềm cho đồng thời khả năng mô phỏng gắn kết và tính toán năng lượng gắn kết chính xác hơn các phần mềm hiện có do đã phối hợp được phương pháp tính nhanh theo nguyên lý thứ nhất (fast

first principle) hay còn gọi là ab-initio với phương pháp cơ học phân tử. Năng lượng gắn kết và IC50 được xác định theo (1). QUANTUM cũng cho giao diện đẹp với các giải pháp hiển thị khác nhau.

Việc thực hiện đột biến gen và tối ưu hình học sản phẩm thu được được thực hiện trên HYPERCHEM phiên bản 7.0 [14]. Trong HYPERCHEM có các giải pháp tính lực khác nhau, hoặc AMBER [15], OPLS [16] và CHARM [17]. Để giảm sai số gây nên do sự khác biệt giữa các kỹ thuật tối ưu hình học khác nhau, chúng tôi đã tối ưu lại cấu hình CD38 tự nhiên trước khi đưa vào các tính toán kế tiếp và sử dụng các tham số lực tương tác OPLS.

Các dữ liệu tính toán gắn kết bằng phần mềm QUANTUM có độ lặp lại rất cao. Sẽ không có thay đổi gì về kết quả giữa các lần tính khác nhau nếu xuất phát từ cùng một bộ dữ liệu ban đầu. Nhưng việc tối ưu cấu hình protein và các thể đột biến ảnh hưởng rất mạnh đến giá trị năng lượng gắn kết phối tử-protein. Tính toán cho thấy, nếu lựa chọn độ hội tụ gradien của phép tính cấu hình tối ưu là 10^{-1} KJ/(A.mol) thì sau khi lặp lại công việc tối ưu (tức là lấy kết quả tối ưu lần trước làm đầu vào), giá trị năng lượng liên kết phối tử-protein thay đổi khoảng 10%(!). Khi giảm độ hội tụ gradien xuống 100 lần thì sự thăng giáng kết quả tính năng lượng gắn kết cũng giảm khoảng 100 lần. Vì thế, những tính toán tối ưu hình học được thực hiện trên HYPERCHEM 7.0 cần đạt tới độ chính xác 10^{-3} KJ/(A.mol).

III - KẾT QUẢ VÀ THẢO LUẬN

1. Đột biến gen và tối ưu hình học

Bảng 1 trình bày kết quả tính toán cấu hình có năng lượng cực tiểu trên phần mềm HYPERCHEM7.0 của CD38 tự nhiên và các thể đột biến E146L, E146F, E146D, E146G, E146A, E146K và E146Q. Các tham số dùng trong quá trình tính toán là: RMS gradien = 0,001 Kcal/(Amol), thuật toán tối ưu là Polak — Ribiere, trường lực cơ học phân tử OPLS, hằng số điện môi = 1, thừa số cân chỉnh tương tác 1-4 đối với lực tĩnh điện là 0,5 và lực Van der Waals là 0,125, không cắt lực xa và tính tất cả các lực

thành phần bao gồm: lực kéo căng, lực cắt, lực xoắn lực van der Waals và lực tĩnh điện. Có thể thấy được rằng tương tác tĩnh điện đóng vai trò rất quan trọng trong việc xác định năng lượng

cơ học phân tử (MM) của protein cũng như cấu hình của nó. Cấu hình có năng lượng cực tiểu thu được của CD38 và các thể đột biến được sử dụng cho các bước tính toán tiếp theo.

Bảng 1: Năng lượng MM cực tiểu của CD38 và các thể đột biến

Thể đột biến	Kéo căng	Cắt kéo	Vặn	Van der Waals	Tĩnh điện
E146F	38,6886	387,120	564,647	-1626,66	-13424,4
E146L	38,6611	385,866	563,029	-1624,64	-13428,7
E146K	38,6356	384,666	577,925	-1623,89	-13398,7
E146G	38,5014	384,370	562,471	-1611,24	-13433,0
E146Q	39,0262	386,828	564,345	-1627,03	-13471,2
E146D	38,6214	386,287	567,294	-1617,73	-13591,3
E146A	38,7098	384,493	562,493	-1614,75	-13433,3
CD38	39,0502	388,302	565,883	-1620,27	-13593,8

Gắn kết phối tử (docking).

Quá trình gắn kết (docking) được thực hiện trên phần mềm QUANTUM 3.3. Thoạt tiên những tính toán gắn kết được thực hiện với các hệ phức protein-phối tử đã có các số liệu thực nghiệm về cấu trúc tinh thể để xác định độ tin cậy của phần mềm. Các hệ được chọn gồm: 2HCT, 2O3Q, 2O3R và 2O3S. Trong số các kết quả đầu ra của QUANTUM 3.3 có: IC50 ($\mu\text{Mol/L}$), E bind (kJ/mol) - năng lượng gắn kết, Ees (kJ/mol) - năng lượng solvat hóa và năng lượng tĩnh điện, Evdw (kJ/mol) - năng

lượng tĩnh điện khoảng cách gần và năng lượng Van der Waals, TdS (kJ/mol)-phần góp của entropy, Etor (kJ/mol)-biến thiên nội năng của phối tử, điện tích Q, khối lượng M, số các liên kết dẻo (flexible) của phối tử n và RMSD - căn của bình phương sai khác trung bình tọa độ các nguyên tử của phối tử giữa giá trị đo được và tính toán. Kết quả tính toán gắn kết phối tử - protein được trình bày trong bảng 2, cùng với kết quả tính bằng phần mềm QSITEFINDER [18] để so sánh.

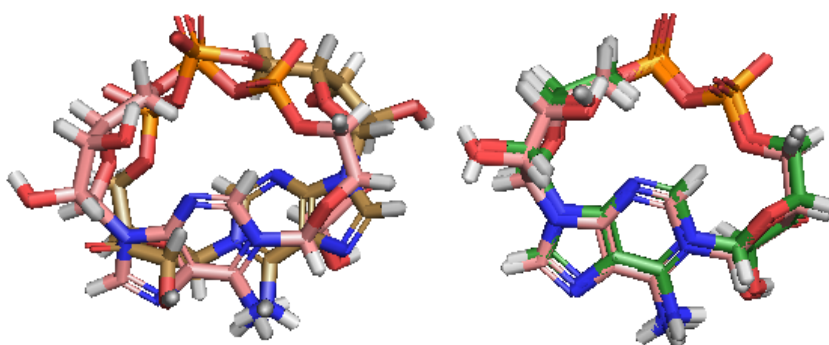
Bảng 2: Kết quả tính toán gắn kết phối tử - protein CD38 và các thể đột biến (2 cấu hình có năng lượng cực tiểu trong môi trường hợp)

STT	E _{gắn kết}	E _{tĩnh điện}	E _{vdw}	TdS	E _{vặn}	RMSD	Qs ⁺
2HCT	-43,2520	-26,6225	-45,8236	-38,3056	-9,1113	0,6842	68
	-32,1916	-10,1466	-40,1298	-22,7869	-4,7021	5,3212	0
2O3Q	-37,0319	-26,7672	-39,9502	-26,4803	3,2051	6,0640	42
	-34,5746	-15,6394	-41,6894	-22,9388	-0,1846	0,2981	0
2O3R	-37,0485	-25,7198	-40,0410	-28,1092	0,6932	0,3198	0
	-34,4071	-25,1102	-37,2226	-24,7035	3,2222	6,1247	50
2O3S	-37,5786	-28,7170	-39,3611	-28,9777	1,5218	6,1885	31
	-36,7304	-26,8606	-38,5805	-29,4607	-0,7500	0,6142	0

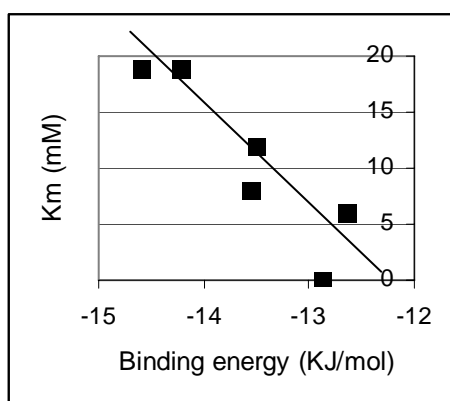
Qs⁺. Độ chính xác của kết quả nhận được từ phần mềm QSITEFINDER [18].

Dường như là năng lượng gắn kết không đồng biến với RMSD. Trong hệ 2O3Q và 2O3S cực tiểu năng lượng gắn kết có RMSD lớn hơn nhiều ($\geq 6,0$) so với cấu hình kế tiếp có năng lượng cao hơn nhưng lại gần như trùng khớp với

dữ liệu tinh thể (RMSD = 0,30 và 0,61). Hình 1 trình bày trường hợp hai cấu hình phức kế tiếp của thể đột biến E226Q của CD38 với cADPR (cyclo-ADPriboz) (mã 2O3Q) nhận được từ QUANTUM.



Hình 1: Vị trí của các phối tử cADPR trong hai cấu hình kế tiếp nhận được từ QUANTUM so sánh với thực nghiệm. Cấu hình a có năng lượng gắn kết lớn hơn cấu hình b nhưng sai khác nhiều hơn so với thực nghiệm



Hình 2: Sự phụ thuộc giữa năng lượng gắn kết NA-protein và hằng số tốc độ phản ứng vòng hóa NAD K_m

Điều này gợi ý việc sử dụng thận trọng hơn QUANTUM khi tính toán mô phỏng các cấu hình phức mà ta không có số liệu thực nghiệm để tham chiếu. Những tính toán dưới đây của chúng tôi chấp nhận cả hai cấu hình kế tiếp có năng lượng gắn kết lớn nhất. Việc loại trừ cấu hình nào cũng đều phải rất thận trọng. Cũng có một cách khác là sử dụng những phần mềm bổ sung. Tuy vậy, không phải lúc nào cũng giúp chúng ta lựa chọn được hợp lý. Cột cuối cùng

của bảng 2 trình bày kết quả tính trực tuyến bằng phần mềm QSITEFINDER [18]. Chỉ có 3 trong số 4 cấu hình được tính toán giúp chúng ta lựa chọn được tốt hơn. Cấu hình thứ nhất của phức 2O3R do QSITEFINDER chọn cũng có RMSD cao hơn hẳn so với cấu hình thứ hai.

Tính toán dự báo năng lượng và cấu hình của các phức phối tử-thể đột biến

Bảng 3 cột 2 trình bày năng lượng gắn kết của NMN với các thể đột biến khác nhau của CD38 ở cấu hình có năng lượng tối ưu so sánh với giá trị tương ứng của CD38 tự nhiên. Có thể dễ dàng thấy được rằng, tất cả các thể đột biến đều có năng lượng gắn kết với phối tử NMN cao hơn hẳn so với dạng tự nhiên CD38. Điều này phù hợp định tính rất tốt với thực nghiệm rằng trong khi CD38 tự nhiên không có hoạt tính xúc tác với quá trình vòng hóa NAD tạo cADPR thì toàn bộ các thể đột biến được khảo sát bởi Graeff và cộng sự [8] đều có hoạt tính này.

Những tính toán tương tự được thực hiện với các phức của thể đột biến với NAD, cADPR và NA (nicotine amide) (xem bảng 3, các cột 3, 5 và 6). Có một sự đồng điệu rõ ràng giữa năng lượng gắn kết của các phức NA-protein với kết quả thực nghiệm. Ngoại trừ trường hợp của thể

đột biến E146G có giá trị năng lượng gắn kết xấp xỉ giá trị của CD38, sự phụ thuộc giữa năng lượng gắn kết NA với các thể đột biến khác của

CD38 và hằng số tốc độ K_m của phản ứng đóng vòng tạo thành cADPR gần như là tuyến tính (hình 2).

Bảng 3: Năng lượng của phức giữa các phối tử NMN, cADPR, NAD và NA với các thể đột biến và hoạt tính xúc tác của CD38 trong phản ứng vòng hóa NAD tạo cADPR (2 cấu hình có năng lượng cực tiểu trong mỗi trường hợp)

Thể đột biến	E _{gắn kết}						Thực nghiệm
	NMN	cADPR	Δ_1	NAD	NA	Δ_2	K_m (#M) ^a
E146L	-28,1434	-28,4007	-0,2573	-38,3173	-13,5028	-3,5862	12□2
	-22,7130	-28,1496		-35,5926	-12,1759		
E146A	-27,9097	-24,8660	3,0437	-33,1435	-14,5932	-6,3157	19□1
	-24,7521	-24,5893		-32,7537	-14,3065		
E146Q	-27,6790				-13,6914		12□2 ^b
	-23,7418				-10,0806		
E146D	-26,3161	-26,9948	-0,6787	-31,7169	-13,5617	-8,8396	8□1
	-22,6381	-25,9828		-28,8530	-13,0348		
E146K	-26,2558				-12,6358		6□6 ^b
	-24,5223				-12,0864		
E146G	-24,1360	-27,4343	-3,2983	-33,9417	-12,6482	-6,1408	6□1
	-23,5075	-26,4949		-30,7606	-12,1312		
E146F	-23,5523	-28,3270	-4,7747	-35,3166	-14,2220	-7,2324	19□2
	-23,3913	-27,3109		-34,2857	-12,2363		
CD38	-31,6989	-27,4104	4,2885	-32,6871	-12,8671	-7,5904	0
	-30,4101	-26,6974		-32,3035	-12,2747		

^aHằng số động học phản ứng vòng hóa [8]. ^bGiá trị dự báo. $\Delta_1 = E_{cADPR} - E_{NMN}$, $\Delta_2 = E_{cADPR} + E_{NA} - E_{NAD}$

Điểm khác biệt cơ bản của các phức với NAD và cADPR so với các phức với NMN và NA là do phân tử lớn, công kênh hơn nhiều nên không thể đi sâu vào vùng các tâm hoạt động - cũng tức là vào vùng lõm (pocket) được xác định bởi các nhóm Glu146 hoặc các nhóm thay thế (bảng 1) và các nhóm Trp125, Trp189 và Glu 226. Hình 3 minh họa cấu trúc có năng lượng cực tiểu của các phức của thể đột biến E146F với cADPR và NA.

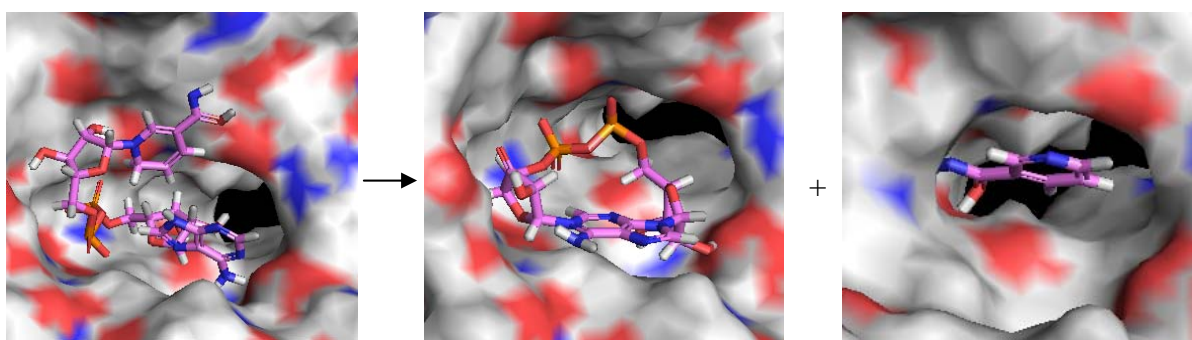
Về cơ chế phản ứng tạo thành cADPR

Như vậy là có tới hai yếu tố tác động đến cơ chế quá trình đóng vòng và hoạt tính xúc tác của CD30 và các thể đột biến: thứ nhất là hình học

và thứ hai là năng lượng. Bảng 3 cột 4 và 7 trình bày sự khác biệt giữa năng lượng gắn kết của các mảnh protein và phối tử tham gia vào phản ứng vòng hóa bao gồm: $\Delta_1 = E_{cADPR} - E_{NMN}$ $\Delta_2 = E_{cADPR} + E_{NA} - E_{NAD}$. Việc tính Δ_1 nhằm xác định vai trò của tương tác protein-NMN trong quá trình vòng hóa NAD. Việc tính Δ_2 nhằm xem xét biến thiên năng lượng của phản ứng $NAD \rightarrow cADPR + NA$. Đáng tiếc là trong cả hai trường hợp đều không thấy rõ mối quan hệ của các biến thiên năng lượng với hằng số động học phản ứng K_m . Cùng với sự phụ thuộc gần tuyến tính của năng lượng gắn kết giữa các thể đột biến với NA và K_m (hình 1) và sự khác biệt rõ ràng về năng

lượng gắn kết NMN với CD38 so sánh với các thể đột biến chúng ta có thể khẳng định được là NA, NMN và năng lượng gắn kết NA-protein, NMN-protein có vai trò quyết định đối với phản ứng vòng hóa NAD. Nếu năng lượng gắn kết

giữa thể đột biến và NA càng lớn thì phản ứng vòng hóa càng thuận lợi. Cơ chế vòng hóa cADPR được minh họa trên hình 3. Cơ chế này phù hợp với cơ chế do Qun Liu và cộng sự đề xuất [10] trên cơ sở thực nghiệm.



Hình 3: Minh họa cấu hình có năng lượng cực tiểu của phức E146F với lần lượt NAD, cADPR, NA và cơ chế phản ứng vòng hóa

Kết quả tính toán của chúng tôi cho thấy, phương pháp tính trình bày ở trên có thể áp dụng hiệu quả và rộng rãi cho việc nghiên cứu tác động đột biến gen đến vai trò xúc tác của các thể đột biến trong các protein khác.

Công trình nhận được sự tài trợ của Bộ Khoa học và công nghệ thông qua đề tài, mã số 5.072.06. Tác giả xin chân thành cảm ơn.

TÀI LIỆU THAM KHẢO

1. Đặng Ứng Vận. Tạp chí Hóa học, T. 45, số 4 (2007).
2. Đặng Ứng Vận. Tạp chí Hóa học, T. 45, số 5 (2007).
3. F. Malavasi, A. Funaro, S. Roggero, A. Horenstein, L. Calosso, and K. Mehta. Immunol. Today, 15, 95 - 97 (1994).
4. E. Ferrero, F. Malavasi. J. Leukocyte Biol., 65, 151 - 161 (1999).
5. H. C. Lee. Biol. Chem. Hoppe-Seyler, 380 785 - 793 (1999).
6. H. C. Lee, C. Munshi, and R. Graeff. Mol. Cell. Biochem., 193, 89 - 98 (1999).
7. H. C. Lee. J. Membr. Biol., 173, 1 - 8 (2000).
8. R. Graeff, C. Munshi, R. Aarhus, M. Johns, and H. C. Lee. J. Biol. Chem., 276, 12169 - 12173 (2001).
9. R. Graeff, Q. Liu, I. A. Kriksunov, Q. Hao and H. C. Lee. J. Biol. Chem., 281 28951 - 28957 (2006).
10. Q. Liu, I. A. Kriksunov, R. Graeff, H. C. Lee and Q. Hao. J. Biol. Chem., 282 5853 - 5861 (2007).
11. QUANTUM 3.3 Docking/Library Screening Software Quantum Pharmaceuticals 2007.
12. GLIDE: Schrodinger, LLC: Portland, OR, 2000.
13. A. E. Cho, V. Guallar, B. J. Berne, R. Friesner. J. Comp. Chem., 26, 915 - 931 (2005).
14. HYPERCHEM™ Release 7.0 for Windows Molecular Modeling System Copyright ©2002 Hypercube, Inc.
15. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. L. Am. Chem. Soc., 117, 5179

- 5197 (1995).
16. W. L. Jorgensen, and J. Tirado-Rives. J. Am. Chem. Soc., 110, 1657 - 1666 (1988).
17. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. J. Kaplus. Comput. Chem., 4, 187 - 21 (1983).
18. A. T. R. Laurie, R. M. Jackson. <http://www.bioinformatics.leeds.ac.uk/qsitefinder>