Published in final edited form as:

Biopolymers. 2009 ; 92(2): 83–93. doi:10.1002/bip.21122.

Influence of the Dye Presence on the Conformational Preferences of CREKA, a Tumor Homing Linear Pentapeptide

David Zanuy¹, David Curcó², Ruth Nussinov^{3,4}, Carlos Alemán¹

¹Departament d'Enginyeria Química, E. T. S. d'Enginyeria Industrial de Barcelona, Universitat Politècnica de Catalunya, Diagonal 647, Barcelona E-08028, Spain

²Departament d'Enginyeria Química, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona E-08028, Spain

³Basic Research Program, SAIC-Frederick, Inc., Center for Cancer Research Nanobiology Program, NCI, Frederick, MD 21702

⁴Department of Human Genetics, Sackler Medical School, Tel Aviv University, Tel Aviv 69978, Israel

Abstract

The conformational properties of the Cys-Arg-Glu-Lys-Ala (CREKA) peptide sequence labeled with fluorescein, a fluorescent dye attached to the Cys through a flexible linker have been examined using molecular dynamics simulations. The CREKA sequence has been identified as a tumor-homing peptide that effectively binds to clotted plasma proteins. Before conformational exploration, the molecular geometry, basicity and spectroscopic properties of this dye, which is essential for the imaging the peptide activity, have been examined using quantum mechanical calculations, with the results also allowing determination of the force-field parameters required for classical simulations. Minimum energy conformations derived from the conformational search have been classified using clustering analyses with criteria based on both the existence of interactions and backbone geometric similarity. The results have been compared with those reported for isolated CREKA (peptide without dye). We found that the fluorescein affects the energy distribution of the minimum energy conformations, with the repulsive steric interactions induced by the dye producing shifting the distribution towards higher energy values.

Interestingly, although the structural characteristics of the bioactive conformation identified for CREKA are not perturbed by the dye, it is less stable when the peptide is attached to the dye than in other chemical environments previously studied (isolated peptide, peptide attached to the surface of a protein, and peptide inserted in a phage display protein).

Keywords

tumor-homing; molecular dynamics; conformation; fluorescein

Correspondence to: David Zanuy; david.zanuy@upc.edu or Carlos Aleman, carlos.aleman@upc.edu.

The content of this publication does not necessarily reflect the view of the policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organization imply endorsement by the US Government.

INTRODUCTION

CREKA (Cys-Arg-Glu-Lys-Ala) is a short linear peptide recently discovered¹ by in vivo screening of phage-display peptide libraries² for tumor-homing in tumor-bearing MMTV-PyMT transgenic breast cancer mice.³ Synthesized CREKA peptide, labeled with the fluorescent dye 5 (6)-carboxyfluorescein (FMA) was detectable in human tumors from minutes to hours after intravenous injection, while it was essentially undetectable in normal tissues (Ruoslahti, 2007). Furthermore, in vivo experiments revealed that CREKA bound to clotted plasma proteins, establishing its being a clot-binding peptide. In addition, experiments with CREKA peptides linked to amino dextran-coated iron oxide (SFIO) nanoparticles indicated that they not only bind to blood and plasma clots but can also effectively induce further localized tumor clotting, thus amplifying the nanoparticle homing.¹ It was further found that the chemical nature of the nanoparticle is not important for this activity since both CREKA-SPIO nanoparticles and CREKA-coated liposomes were found to cause clotting in tumor vessels.¹

In a recent study, we determined the energy landscape and the bioactive conformation of CREKA using a multiple conformational search strategy based on molecular dynamics (MD) simulations.⁴ We considered two experimentally tested environments: (i) the free peptide, which was immersed in both a dilute aqueous solution and an ionized aqueous solution that mimics the physiological medium; and (ii) the peptide attached to a nanoparticle, in which the sulfhydryl group of the Cys residue was attached to the center of a square surface formed by 100 spherical particles (10×10). In addition, we also considered a molecular system that mimics the CREKA peptide when it is inserted in a phage display protein, even though in this case the conformational space was not explored exhaustively, i.e., the search strategy did not completely converge. Results indicated that the conformational profile of the REKA sequence is very similar in all cases, with the only difference being the Cys residue. Complete clustering analyses based on the number and distribution of intramolecular interactions showed that such coincidence was particularly remarkable for the accessible minima, i.e., those with relative energies lower than 2.0 kcal/mol. Furthermore, the coincidence between the global minimum of the studied systems together with the analysis of the accessible conformations allowed us to propose that the bioactive conformation of CREKA presents both a β -turn motif and strong interactions involving the side chains of Arg, Glu, and Lys. The MD simulations further showed that this was a very stable conformation.

In our previous study, we did not consider the experimentally tested system formed by the CREKA peptide labeled with the fluorescent dye FMA, which are connected through a linker that involve five methylene units. However, a detailed inspection of this system, hereafter denoted FMA-CREKA (see Figure 1), suggests that its energy landscape may be significantly different from that found for the free peptide. Thus, the polar groups of the fluorescent dye, which could be easily accessed by the peptide due to the substantial conformational flexibility of the linker, may induce significant changes in the interaction pattern of the peptide. Accordingly, comparison between the energy landscapes of the free peptide and FMA-CREKA may lead to a re-definition of the bioactive conformation; or,

alternatively, to a validation of the conformational profile of CREKA independent of the chemical environment.

In this work, we have used classical MD simulations to explore the conformational properties of FMA-CREKA in aqueous solution, which have been compared with those reported for isolated CREKA⁴ Before undertaking the conformational search of the peptide covalently attached to the fluorescent dye, the properties of FMA have been examined using quantum mechanical calculations. The results, which have allowed us to determine the conformation, molecular geometry, basicity and spectroscopic properties of the dye, have also been used to develop the force-field parameters necessary for MD simulations. Finally, the bioactive conformation of FMA-CREKA is discussed and compared with that of CREKA, examining the influence of the dye on the structural characteristics and energy of this conformation.

METHODS

Quantum Mechanical Calculations

The molecular geometries of the neutral and ionized forms of FMA were optimized using the restricted and unrestricted formalisms, respectively, of the B3LYP^{5,6} method combined with the 6-31+G(d,p) basis set.⁷ The dye was blocked with the *N*-methylacetamide group to mimic the linker used in FMA-CREKA (see Figure 1).

To obtain an estimation of the solvation effects on the relative stability and the properties of the dye, single point calculations were conducted on the optimized structures using a SCRF model. Specifically, the polarizable continuum model (PCM) developed by Tomasi and coworkers⁸⁻¹⁰ was used to describe water as solvent. The PCM model represents the polarization of the liquid by a charge density appearing on the surface of the cavity created in the solvent. This cavity is built using a molecular shape algorithm. PCM calculations were performed in the framework of the (U)B3LYP/6-31+G(d,p) level using the standard protocol and considering the dielectric constants of water ($\epsilon = 78.4$) to obtain the free energies of solvation (G_{sol}) of the minimum energy conformations. Within this context, it should be emphasized that previous studies indicated that solute geometry relaxations in solution and single point calculations on the optimized geometries in the gas phase give almost identical G_{solv} values.^{11,12}

All quantum mechanical calculations were carried out using the Gaussian 03 computer program.¹³

Conformational Search and MD Simulations

The conformational preferences of FMA-CREKA were explored using the same procedure as in our previous work,⁴ which is based on a modified simulated annealing procedure combined with MD (SA-MD) as a sampling technique.^{14,15} More specifically, our strategy consists of performing independent SA-MD cycles and selection of a large number of structures (500) from each SA-MD cycle for energy minimization. This procedure is robust enough to locate the lower-energy minimum structures of the system under study, i.e., structures that are quasi-degenerate with the global minimum but situated in different valleys

of the peptide landscape. The development of this sampling technique was inspired by recent studies demonstrating that very low energies are obtained by minimizing the energy of structures generated at the initial and intermediate states of conventional SA-MD, although application of the latter procedure does not always lead the system to the most stable region at the end of the process.^{16,17}

The simulated peptide was placed in the center of a cubic simulation box ($a = 47.5 \text{ \AA}$) filled with at least 3405 explicit water molecules, which were represented using the TIP3 model.¹⁸ Two negatively charged chloride and one positively charged sodium atoms were added to the simulation box to reach electric neutrality (net charges were considered for Arg, Lys, and Glu residues at neutral pH). The energy was calculated using the AMBER force-field.^{19,20} All parameters were taken from the AMBER libraries with exception of the partial atomic charges for the dye and the linker. The electrostatic parameters were obtained by fitting the molecular electrostatic potential calculated at the *ab initio* HF/6–31G(d)²¹ level to that obtained using atom centered partial charges. It should be noted that this methodology provides electrostatic parameters compatible with the AMBER force-field.¹⁹ Atom pair distance cut-offs were applied at $14.0 \text{ \AA}$ to compute the van der Waals and electrostatic interactions. Both temperature and pressure were controlled by the weak coupling method, the Berendsen thermobarostat,²² using a time constant for heat bath coupling and a pressure relaxation time of 1 ps. Bond lengths were constrained using the SHAKE algorithm²³ with a numerical integration step of 2 fs. All classical calculations were performed using the NAMD computer program.²⁴

The simulation box was equilibrated before the conformational search. For this purpose, 0.5 ns of NVT-MD at 500 K were used to homogeneously distribute the solvent and ions in the box. Next, 0.5 ns of NVT-MD at 298 K (thermal equilibration) and 0.5 ns of NPT-MD at 298 K (density relaxation) were run. The last snapshot of the NPT-MD was used as starting point for the conformational search process. This initial structure was quickly heated to 900 K at a rate of 50 K/ps to force the molecule to jump to a different region of the conformational space. Along 10 ns, the 900 K-structure was slowly cooled to 500 K at a rate of 1 K per 25 ps. A total of 500 structures were selected and subsequently minimized during the first cycle of modified SA-MD. The resulting minimum energy conformations were stored in a rank-ordered library of low energy structures. The lowest energy minimum generated in a modified SA-MD cycle was used as starting conformation of the next cycle.

Conformation Classification and Clustering Analysis

To construct a list of unique minimum energy conformations, the 500 minimized structures provided by each cycle of modified SA-MD were compared not only among themselves but also with the unique minima generated in previous cycles. The list was organized by rank ordering all the unique minimum energy conformations found following an increasing order of energy. Previously listed conformations were discarded. The criterion used to identify unique minimum energy conformations was based on the values of virtual dihedral angles, which were defined using main-chain atoms, and the presence of interaction patterns, i.e., salt bridges, hydrogen bonds, and dipole-dipole interactions.

Five virtual dihedral angles were defined considering the α -carbon atoms of the peptide within the FMA-CREKA system (see Figure 1). Each of these virtual dihedral angles was associated with the conformation of one amino acid of the peptide. Interactions were recorded based on the following geometric criteria: (a) salt bridges: the distance between the centers of the interacting groups shorter than 4.50 Å; and (b) hydrogen bonds: distance $H \cdots O$ ($d_{H \cdots O}$) shorter than 2.50 Å and angle $\angle N - H \cdots O$ higher than 120.0°.

Two structures were considered different if they differ in at least one of their virtual dihedral angles by more than 60° or in at least one of the interactions counted.

RESULTS AND DISCUSSION

Properties of the FMA

To provide a complete characterization of the dye, the molecular geometries of the neutral, monoanionic, and dianionic species of the dye, i.e., FMA(0), FMA(−1), and FMA(−2), respectively, were optimized at the (U)B3LYP/6–31+G(d,p) level. The chemical structures of these species, which are displayed in Scheme 1, show that the monocationic and dicationic forms were generated by deprotonation of the hydroxyl groups attached to rings A and C (see Scheme 1), respectively.

Geometry optimizations reveal that the two ring systems adopt a perpendicular arrangement in all cases, this feature being illustrated in Figure 2 for FMA(0). Indeed, although the ionization state does not produce significant changes in the relative disposition of the two molecular fragments involving fused rings, it alters considerably the bond lengths of the rings labeled as A, B, and C (see Figure 1). This is observed in Table I, which compares the bond lengths obtained for FMA(0), FMA(−1), and FMA(−2). As can be expected, the largest effect is detected for the ring containing the deprotonated hydroxyl groups (rings A and C), even though the C2-O and C9-O bond lengths also underwent a significant change upon the deprotonation of the hydroxyl groups. On the other hand, ionization does not alter significantly the bond lengths of the rings D and E (data not shown).

The proton affinity (PA), which describes the basicity in the gas-phase of a given base, was estimated for FMA(0) and FMA(−1) using the following approach:

$$PA[FMA(0)] \approx \Delta E_{\text{ele}}[FMA(0) - FMA(-1)] \quad (1)$$

$$PA[FMA(-1)] \approx \Delta E_{\text{ele}}[FMA(-1) - FMA(-2)] \quad (2)$$

where E_{ele} corresponds to the difference between the electronic energies at 0 K. Thus, due to the size of the system under study, statistical corrections were neglected. However, previous studies in chemically related systems, i.e., hydroxylated flavonoids,^{25,26} indicated that the approach used in this study provides a good qualitative description of the basicity in the gas-phase, zero-point energies, entropic and thermal corrections being only useful for quantitative descriptions, i.e., statistical corrections produces a systematic reductions of the resulting PAs.

The PA estimated for FMA(0) and EMA(-1) is 336 and 375 kcal/mol, respectively, indicating that, as expected, the former is more acidic than the latter. However, the most striking feature is that the PAs are very large. Thus, PAs of benzopyran-4-ones, which also involve one hydroxyl group, ranged from 256 to 279 kcal/mol depending on the position of the acidic substituent.²⁵ It should be noted that such values were estimated using the same approach, i.e., neglecting zero-point energies, entropical and thermal corrections, and a similar theoretical level, i.e., (U)MP2/6-31+G(d). On the other hand, as expected the strength of the interaction between FMA and the aqueous solvent increases with the charge of the solute, G_{sol} being -23.3, -63.1, and -149.5 kcal/mol for FMA(0), FMA(-1), and FMA(-2), respectively.

The relative pK_a (pK_a) between the FMA(0) and FMA(-1) was estimated according to²⁷:

$$\Delta pK_a = \frac{1}{2.303RT} (\text{PA}[\text{FMA}(0)] - \text{PA}[\text{FMA}(-1)] + \Delta\Delta G_{\text{sol}}^{\text{np}}) \quad (3)$$

where $\Delta\Delta G_{\text{sol}}^{\text{np}}$ is the difference between free energies of solvation (G_{sol}) for non-protonated and protonated forms.

$$\begin{aligned} \Delta\Delta G_{\text{sol}}^{\text{np}} &= \{ \Delta G_{\text{sol}}[\text{FMA}(-1)] - \Delta G_{\text{sol}}[\text{FMA}(0)] \} \\ &\quad - \{ \Delta G_{\text{sol}}[\text{FMA}(-2)] - \Delta G_{\text{sol}}[\text{FMA}(-1)] \} \\ &= 2\Delta G_{\text{sol}}[\text{FMA}(-1)] - \Delta G_{\text{sol}}[\text{FMA}(0)] \\ &\quad - \Delta G_{\text{sol}}[\text{FMA}(-2)] \end{aligned} \quad (4)$$

The resulting pK_a is 7, this value being slightly larger than that expected for independent hydroxyl groups ($pK_a \sim 17$) with similar chemical environments. However, the deviation between the calculated and experimental pK_a s should be attributed to both the approximations used to estimate PAs, i.e., omission of thermodynamic corrections, and to the limitations of the SCRF solvation models, i.e., omission of specific interaction between the solute and the first solvation shell. However, in spite of this deviation, the present results indicate that deprotonation of hydroxyl groups in FMA are not coupled to electronic effects, with the processes being independent.

On the basis of the calculated pK_a values, the neutral dye species was used for MD simulations of FMA-CREKA. The molecular geometry of FMA(0) was used to define the equilibrium geometry parameters (bond lengths and bond angles), whereas partial electrostatic charges were developed using the procedure indicated in the Methods section. Figure 3 displays the electrostatic parameters used for the dye and the linker in the conformational exploration presented in the next section.

Energy Landscape of FMA-CREKA

The evolution of the number of unique minimum energy conformations against the number of modified SA-MD cycles for both FAM-CREKA and CREKA are compared in Figure 4a, whereas the energy distributions of the unique minimum energy conformations identified for these are shown in Figure 4b. As can be seen, the conformational search of the two peptides converges after 13 production cycles, the number of unique minimum energy conformations

being 3188 and 3130 for FMA-CREKA and CREKA, respectively. However, although the dye does not affect the efficacy of the sampling procedure, it is noticeable that it influences the energy distribution of the structures. More specifically, the increment in the steric hindrance shifts the global distribution of FMA-CREKA towards higher energy values, the average increment ranging from 4 to 8 kcal/mol. This is not a dramatic change, even though it could induce significant variations in the bioactive conformation.

Figure 5a represents the distribution of virtual dihedral angles used to define the conformational preferences of the five residues in FMA-CREKA. Comparison of this distribution with that obtained for isolated CREKA, which was previously reported,⁴ indicates that Cys is the only residue affected by the covalent attachment with the dye. This is observed in Figure 5b, which compares the distribution of the main chain dihedral angles ϕ , ψ for the significant (accessible) unique minima of FMA-CREKA and CREKA, i.e., the 23 and 27 minima with relative energy lower than 2.0 kcal/mol, respectively. The results obtained for the two systems show a substantial resemblance, even though the attachment of the dye allows the ϕ angle of the Cys residue to explore wider regions on the Ramachandran surface. However, the impact of the latter event does not alter the conformational preferences of the other residues. This is consistent with experimental observations that show similar recognition efficiency in both CREKA and FMA-CREKA.

Clustering analyses based on the existence of intramolecular hydrogen bonds and salt-bridges indicate that 75.6% (2410 structures) and 69.2% (2165 structures) of the minima obtained for FMA-CREKA and CREKA, respectively, present at least one of such interactions. The total number of interactions computed for such structures reaches 4627 and 4243 for FMA-CREKA and CREKA, respectively, the average number of interactions per structure being about 1.9 in both cases. The different combination of interactions leads to a distribution of these structures in 683 and 515 clusters, respectively, even though only a small number of such dusters contain a significant number of structures. Specifically, the number of clusters that involve at least 25 unique minimum energy structures was 15 and 13 for FMA-CREKA and CREKA, respectively. This is reflected in Figure 6, which represents the number of minimum energy structures contained in each cluster for both FMA-CREKA and CREKA. A detailed analysis of these structures reveals that the interactions are distributed as follows: (i) main chain · · · main chain hydrogen bonds involving two backbone amide groups: 85.7% (FMA-CREKA) and 84.7% (CREKA); (ii) main chain · · · side chain hydrogen bonds between one backbone amide group and the side chain of Arg: 6.7% (FMA-CREKA) and 6.7% (CREKA); and (iii) side chain · · · side chain salt bridges: 7.4% (FMA-CREKA) and 8.6% (CREKA). The most frequent interactions of each category found in both FMA-CREKA and CREKA are detailed in Table II. As can be seen, main chain · · · main chain interactions involve a large number of intramolecular hydrogen bond types, whereas side chain · · · side chain hydrogen bonds and side chain · · · side chain salt bridges are concentrated in a few types of interactions. However, the most relevant feature in Table II is that the interaction patterns of FMA-CREKA and CREKA are very similar, the largest difference being 3%.

A complementary clustering analysis based on geometrical considerations rather than on interactions was developed to compare the conformational preferences of FMA-CREKA and

CREKA. Specifically, the root mean square deviation (RMSD) between the backbone atoms was used to categorize the minimum energy structures in clusters. Within each duster the RMSD of all the structures contained in it, which is calculated with respect to the lowest energy one, is smaller than 1.50 Å. The choice of this threshold is based on the fact that the oscillations produced in the main chain RMSD of stable conformations by the thermal agitation is lower than 1.50 Å, i.e., the variation of the RMSD is typically around 1.2 Å.⁴ Obviously, the dye and the linker of FMA-CREKA structures were omitted for the evaluation of the corresponding RMSDs. This geometric criterion produced a total of 37 and 35 dusters from a total of 3188 (FMA-CREKA) and 3130 (CREKA) minima, respectively. The results are depicted in Figure 7.

Interestingly, the distribution of clusters obtained for FMA-CREKA differs significantly from that of CREKA. Indeed, a detailed analysis of the minimum energy structures indicates that the more important conformational differences appear in the main chain atoms. Thus, the number of minima contained in the different dusters of the two systems is significantly different, especially when dusters located at the regions of lower energy are compared. In spite of this difference, a detailed comparison of the lowest energy structure, hereafter denoted leading structure, of each duster for the two systems under study reveals important structural and energetic features.

Figure 8 compares the leading structures of the more stable dusters of CREKA with the leading structure of the FMA-CREKA dusters that shows highest conformational similarity, the superposition between the structures being also displayed to illustrate the distortions induced by the dye in the peptide. As can be seen, most of the dusters found for FMA-CREKA share conformational features with those obtained for CREKA, even though there are differences in the relative energy order of the two systems. Three of the top four CREKA clusters are also present in the top five FMA-CREKA clusters. Moreover, the lowest energy structure of CREKA (Figure 8a), which was previously identified as the bioactive conformation because of its unaltered stability in different chemical environments,⁴ is very similar to the leading structure of the third duster found for FMA-CREKA. Overall, these results indicate that the interactions between the dye and the peptide side chains are responsible for some alterations in the relative energy order of the dusters. For example, the leading structure of the second FMA-CREKA cluster is significantly stabilized by the interaction between the dye and the side chain of Lys (Figure 8e). This structure coincides with the leading structure of the 24th CREKA duster, which is much less stable because the formation of such interaction is unlikely.

The clustering analyses of the minimum energy conformations obtained for FMA-CREKA allow us to conclude that the presence of the dye and the linker do not alter the structural characteristics of the bioactive conformation previously identified for CREKA,⁴ which was found by comparing the minima of this peptide when it is in the isolated state (diluted aqueous solution), attached to the surface of a nanoparticle and inserted in a phage-display protein. The more important characteristics of this bioactive conformation, which is displayed in Figure 8a for both FMA-CREKA (central panel) and CREKA (left panel), are (i) the β -turn motif that involves the Cys and Lys residues; and (ii) the strong interactions among the side chains of Arg, Glu, and Lys. As evidenced in Figure 8a, the shape of the

bioactive conformation looks like a pocket with the charged groups pointing outward, making it optimal for formation of intermolecular interactions with a receptor. Thus, the arrangements of both the peptide backbone and side chain groups are practically identical for both FMA-CREKA and CREKA (see the superposition in the right panel of Figure 8a); however, the interactions of the dye with the peptide backbone reduce the stability of the bioactive conformation. Thus, although the bioactive conformation corresponds to the global minimum in CREKA, this FMA-CREKA minimum is destabilized with respect to the lowest energy one by 0.8 kcal/mol. However, this destabilization does not present a problem; it is well-known that the bioactive conformation does not necessarily correspond to the lowest minimum energy conformation, even though it must be contained within the set of accessible energy minima, i.e., the energy of the bioactive conformation is relatively close to that of the global minimum.

CONCLUSIONS

The molecular geometry, conformation, basicity, and transition energies of the neutral and ionized forms of FMA have been characterized using DFT calculations. In addition, atomic electrostatic parameters, which have been employed in the classical MD simulations of FMA-CREKA, have been developed for the neutral form of this dye. The conformational preferences of FMA-CREKA have been explored using modified SA-MD procedure, which was previously used to investigate the bioactive conformation of the CREKA peptide in different chemical environments: the free state, attached to the surface of a nanoparticle and inserted in a phage display protein.

A total of 3188 unique minimum energy conformations were found in the 13 cycles of SA-MD required to converge the conformational search process. All these minima were classified according to two complementary clustering analyses, with the criteria based on the existence of interactions and geometric similarity of the backbone, respectively. Comparison with the results obtained with the isolated CREKA peptide indicates that, interestingly, the FMA dye alters the energy distribution of the minimum energy conformations. Specifically, the repulsive steric interactions induced by the dye leads to a shift towards higher energy values. Yet, in spite of this, the conformational profiles of FMA-CREKA and CREKA are relatively similar, with the largest difference being the relative energy order of the leading structures in the clusters rather than conformational changes.

On the other hand, the bioactive conformation proposed for the CREKA sequence, which was characterized by a β -turn motif and strong interactions among the side chains of the charged residues, was also detected within the set of low-energy minima of FMA-CREKA. Thus, the bioactive conformation was 0.8 kcal/mol unfavored with respect to the global minimum in FMA-CREKA, even though the bioactive conformation was identified as the lowest energy minimum when the dye is not attached to the peptide, i.e., the bioactive conformation was the global minimum for the CREKA peptide in the free state, attached to a nanoparticle and inserted in a phage display protein. The overall conclusion from the results presented in this work is fully consistent with the experimentally determined activity of this tumor homing peptide when it is used in different chemical environments.

Acknowledgments

The authors are indebted to both the Barcelona Supercomputer Center (BSC) and the “Centre de Supercomputació de Catalunya” (CESCA) for computational resources. They acknowledge the National Cancer Institute for partial allocation of computing time and staff support at the Advanced Biomedical Computing Center of the Frederick Cancer Research and Development Center. Classic calculations were partially formed by utilizing the high-performance computational capabilities of the Biowulf PC/Linux cluster at the National Institutes of Health, Bethesda, MD (<http://biowulf.nih.gov>).

Contract grant sponsor: National Cancer Institute, National Institutes of Health

Contract grant number: N01-CO-12400

Contract grant sponsors: Spanish “*Ministerio de Educación y Ciencia*” (Ramon y Cajal program)

REFERENCES

1. Simberg D; Duza T; Park JH; Essier M; Pilch J; Zhang L; Derfus AM; Yang M; Hoffman RM; Bathia S; Sailor MJ; Ruoslahti E Proc Natl Acad Sci USA 2007, 104, 932–936. [PubMed: 17215365]
2. (a). Hoffman JA; Giraudo E; Singh M; Zhang L; Inoue M; Porkka K; Hanahan D; Ruoslahti E Cancer Cell 2003, 4, 383–391; [PubMed: 14667505] (b) Pasqualini R; Ruoslahti E Nature 1996, 380, 364–366. [PubMed: 8598934]
3. Hutchinson JN; Muller WJ Oncogene 2000, 19, 6130–6137. [PubMed: 11156526]
4. Zanuy D; Flores-Ortega A; Casanovas J; Curcó D; Nussinov R; Alemán C J Phys Chem B 2008, 112, 8692–8700. [PubMed: 18588341]
5. Becke AD J Chem Phys 1993, 98, 1372–1377.
6. Lee C; Yang W; Parr RG Phys Rev B 1988, 37, 785–789.
7. McLean AD; Chandler GS J Chem Phys 1980, 72, 5639–5648.
8. Miertus S; Scrocco E; Tomasi J Chem Phys 1981, 55, 117–129.
9. (a). Miertus S; Tomasi J Chem Phys 1982, 65, 239–245; (b) Tomasi J; Persico M Chem Phys 1994, 94, 2027–2094.
10. Tomasi J; Mennucci B; Cammi R Chem Rev 2005, 105, 2999–3093. [PubMed: 16092826]
11. Jang YH; Goddard WA III; Noyes KT; Sowers LC; Hwang S; Chung DS J Phys Chem B 2003, 107, 344–357.
12. Iribarren JL; Casanovas J; Zanuy D; Alemán C Chem Phys 2004, 302, 77–83.
13. Gaussian 03, Revision B. 02, Frisch MJ; Trucks GW; Schlegel HB; Scuseria GE; Robb MA; Cheeseman JR; Montgomery JA; Vreven T Jr.; Kudin KN; Burant JC; Millam JM; Iyengar SS; Tomasi J; Barone V; Mennucci B; Cossi M; Scalmani G; Rega N; Petersson GA; Nakatsuji H; Hada M; Ehara M; Toyota K; Fukuda R; Hasegawa J; Ishida M; Nakajima T; Honda Y; Kitao O; Nakai H; Kiene M; Li X; Knox JE; Hratchian HP; Cross JB; Adamo C; Jaramillo J; Gomperts R; Stratmann RE; Yazyev O; Austin AJ; Cammi R; Pomelli C; Ochterski JW; Ayala PY; Moikuma K; Voth GA; Salvador P; Dannenberg JJ; Zakrzewski VG; Dapprich S; Daniels AD; Strain C, M.; Farkas O; Malick DK; Rabuck AD; Raghavachari K; Foresman JB; Ortiz JV; Cui Q; Baboul AG; Clifford S; Cioslowski J; Stefanov BB; Liu G; Liashenko A; Piskorz P; Komaromi I; Martin RL; Fox DJ; Keith T; Al-Laham MA; Peng CY; Nanayakkara A; Challacombe M; Gill PMW; Johnson B; Chen W; Wong MW; Gonzalez C; Pople JA Gaussian: Pittsburgh, PA, 2003.
14. Kirkpatrick S; Gelatt CD Jr.; Vecchi MP Science 1983, 220, 671–680. [PubMed: 17813860]
15. Steinbach PJ; Brooks BR Chem Phys Lett 1994, 226, 447–452.
16. Baysal C; Meirovitch HJ Comput Chem 1999, 20, 1659–1670.
17. Simmerling C; Elber R J Am Chem Soc 1994, 116, 2534r–2547.
18. Jorgensen WL; Chandrasekhar J; Madura JD; Impey RW; Klein ML J Chem Phys 1983, 79, 926–935.
19. Wang J; Cieplak P; Kollman PA J Comput Chem 2000, 21, 1049–1074.

20. Cornell WD; Cieplak P; Bayly CI; Gould IR; Merz KM; Ferguson DM; Spellmeyer DC; Fox T; Caldwell JW; Kollman PA J Am Chem Soc 1995, 117, 5179–5197.
21. Hariharan PC; Pople JA Theor Chim Acta 1973, 28, 213–222.
22. Berendsen HJC; Postma JPM; van Gunsteren WE; DiNola A; Haak JR J Chem Phys 1984, 81, 3684–3690.
23. Ryckaert JR; Ciccotti G; Berendsen HJC J Comput Phys 1977, 23, 327–341.
24. Phillips JC; Braun R; Wang W; Gumbart J; Tajkhorshid E; Villa E; Chipot C; Skeel RD; Kale L; Schulten KJ Comput Chem 2005, 26, 1781–1802.
25. Alemán C; Ishiki HM J Org Chem 1999, 64, 1768–1769. [PubMed: 11674257]
26. Alemán C J Mol Struct (Theochem) 2000, 528, 65–73.
27. Lim C; Bashford D; Karplus M J Phys Chem 1991, 95, 5610–5620.

FIGURE 1.

Chemical structure of the dye, linker and peptide in FMA-CREKA. Black dots indicate the atoms used to define the virtual dihedral angles considered in the clustering analysis. The five rings of FMA have been labeled with letters for identification. Atoms of rings A, B, and C have been numbered.

FIGURE 2. Minimum energy conformation of FMA(0) calculated at the B3LYF/6-31+G(d,p) level. The perpendicular arrangement of the two ring systems contained in FMA is clearly illustrated.

Biopolymers. Author manuscript; available in PMC 2020 May 22.

FIGURE 4.

Number of unique minimum energy conformations against the number of modified SA-MD cycles used in the conformational search process (a) and distribution of energies for the unique minimum energy conformations found in the conformational search process (b). In the latter plots the energies are computed relative to the corresponding global minimum. In both cases, the gray line with diamonds and the black line with circles correspond to FMA-CREKA and CREKA systems, respectively.

FIGURE 5.

(a) Distribution of the virtual dihedral angles (see Methods section) used to define the conformation of five residues in all the unique minima obtained for FMA-CREKA. Color code for the bars is: black for dihedral angle values ranging from 0° to 60°, blue from 60° to 120°, red from 120° to 180°, green from 180° to 240°, dark gray from 240° to 300°, and light gray from 300° to 360°. (b) Distribution of the main chain dihedral angles ϕ , ψ for the five residues of FMA-CREKA (gray diamonds) and isolated CREKA (black circles) in the accessible unique minima, i.e., those with relative energy lower than 2.0 kcal/mol.

FIGURE 6. Number of structures in each of the 683 and 515 dusters found for FMA-CREKA (gray diamonds) and CREKA (black circles), respectively. The dustering analysis was based on the existence of interactions, i.e., hydrogen bonds and salt bridges.

FIGURE 7.

Number of structures in each of the 37 and 35 clusters found for 6-FMA/CREKA (top) and isolated CREKA (bottom), respectively. The clustering analysis was based on a geometric criterion, i.e., hydrogen bonds and salt bridges. Within each cluster the RMSD of all the structures contained in it, which is calculated with respect to the lowest energy one, is smaller than 1.50 Å. Clusters are ordered considering the energy of the lead structures (from the lowest to highest), which corresponds the lowest energy structure within the cluster. In 6-FMA/CREKA structures, the dye and the linker have been omitted to compute the RMSD.

FIGURE 8.

Detailed representation of the leading structure of the more stable clusters found for CREKA (left), the leading structure of the FMA-CREKA cluster with highest conformational similarity (centre), and superposition of the two leading structures (right). The relative order, in terms of energies, is indicated in each case for both CREKA and FMA-CREKA clusters. In the individual representation of the leading structures, the peptide backbone has been highlighted using solid tubes, whereas the heavy atoms of both side chains and the dye have

been described using solid lines. Peptide atoms follow the CPK color convention. The dye and the linker in FMA-CREKA are colored in green.

SCHEME 1.
Acid-base equilibrium of FMA.

Table IBond Lengths^a (in Å) Obtained for the Lowest Energy Conformation of FMA(0), FMA(−1), and FMA(−2)

Bond Length	FMA(0)	FMA(−1)-I	FMA(−2)
C1-C2	1.396	1.379	1.377
C2-C3	1.401	1.411	1.432
C3-C4	1.404	1.422	1.431
C4-C5	1.388	1.375	1.368
C5-C6	1.405	1.457	1.462
C6-C1	1.391	1.447	1.445
C6-O	1.366	1.263	1.264
C2-O	1.371	1.397	1.374
O-C9	1.369	1.352	1.373
C8-C9	1.401	1.407	1.431
C7-C8	1.512	1.505	1.413
C3-C7	1.512	1.475	1.431
C8-C13	1.404	1.402	1.431
C13-C12	1.389	1.391	1.369
C11-C12	1.405	1.403	1.461
C10-C11	1.391	1.389	1.445
C9-C10	1.397	1.402	1.378
C11-O	1.366	1.377	1.214

Molecular geometries were obtained at the (U)B3LYP/6–31+G(d,p) level.

^aThe numbering of the molecule is displayed in Figure 1.

Table II

Clustering of FMA-CREKA and CREKA Based on Hydrogen Bonds and Salt Bridges

	FMA/CREKA (%)	CREKA (%)
(i) Main chain . . . main chain hydrogen bonds		
Lys (N-H) . . . (O=C) Cys	9.6	11.5
Ala (N-H) . . . (O=C) Arg	7.0	9.5
Glu(N-H) . . . (O=C) Ace ^a	10.4	9.1
Arg (N-H) . . . (O=C) Ace ^a	11.1	8.1
Ala (N-H) . . . (O=C) Glu	7.4	7.0
NMe (N-H) . . . (O=C) Glu	7.1	6.7
Lys (N-H) . . . (O=C) Ace ^a	6.1	6.3
Lys (N-H) . . . (O=C) Arg	4.4	4.8
Ala (N-H) . . . (O=C) Cys	4.3	4.8
Glu (N-H) . . . (O=C) Cys	5.0	4.3
NMe (N-H) . . . (O=C) Cys	3.7	4.3
(ii) Main chain . . . side chain hydrogen bonds		
Arg side chain . . . (O=C) Ace ^a	3.4	4.7
(iii) Side chain . . . side chain salt bridges		
Arg . . . Glu	5.4	4.3
Lys . . . Glu	2.0	4.3

Within each category (see text), only the interactions with a population higher than 4% for one of the two systems are displayed.

^a Ace refers to the terminal acetyl group of CREKA peptide, which is replaced by the linker in FMA-CREKA (see Figure 1).