Finding the transition state without initial guess: the growing string method for Newton trajectory to isomerisation and enantiomerisation reaction of alanine dipeptide and poly(15)alanine

WOLFGANG QUAPP[a]

[a]Mathematical Institute, University of Leipzig,
Augustus-Platz, D-04109 Leipzig, Germany
E-mail: quapp@rz.uni-leipzig.de

December 18, 2006

Telephone: [49] 341-97 32153
Fax: [49] 341-97 32199
Web: www.math.uni-leipzig.de/∼quapp

Revision, submitted to Journal of Computational Chemistry

Abstract: We report a new, high-dimensional application of a method for finding a transition state (TS) between a reactant and a product on the potential energy surface: the search of a growing string along a reaction path defined by any Newton trajectory in combination with the Berny method (W.Quapp, J. Chem. Phys. 122 (2005) 174106; we have provided this algorithm on a web page). Two given minima are connected by a one-dimensional, but usually curvilinear reaction coordinate. It leads to the TS region. The application of the method to alanine dipeptide finds the TS of the isomerisation $C_{7\,ax} \rightarrow C_5$, some TSs of the enantiomerisation of $C_{7\,ax}$ from L-form to quasi-D-form, and it finds the TS region of a transition of a partly unfolded, bent structure which turns back into a mainly α-helix in the $Ac(Ala)_{15}NHMe$ polyalanine (all at the quantum mechanical level B3LYP/6-31G: the growing string calculation is interfaced with the Gaussian03 package). The formation or dissolvation of some α- or 3_{10}-hydrogen bonds of the helix are discussed along the TS pathway, as well as the case of an enantiomer at the central residue of the helix.

Keywords: Potential energy surface; saddle point; reaction path; Newton trajectory; projection; alanine dipeptide; polyalanine; 2_7-, 3_{10}-, α-H-bond;

Subject area: New application of theoretical methods and algorithms

Graphical Abstract: A high-dimensional application is given of a method for finding a transition state (TS) along the reaction pathway between a reactant and a product on the potential energy surface: the search of a growing string defined by any Newton trajectory (W.Quapp, J. Chem. Phys. 122 (2005) 174106). The application of the method to alanine dipeptide finds the TS of the isomerisation $C_{7ax} \rightarrow C_5$, some TSs of the enantiomerisation of C_{7ax} from L-form to quasi-D-form (given in the graphics), and it finds the TS region of a transition of a partly unfolded, bent structure which turns back into a mainly α-helix in the $Ac(Ala)_{15}NHMe$ polyalanine. The formation or dissolvation of some α- or 3_{10}-hydrogen bonds of the helix are discussed along the TS pathway, as well as the case of an enantiomer at the central residue of the helix.

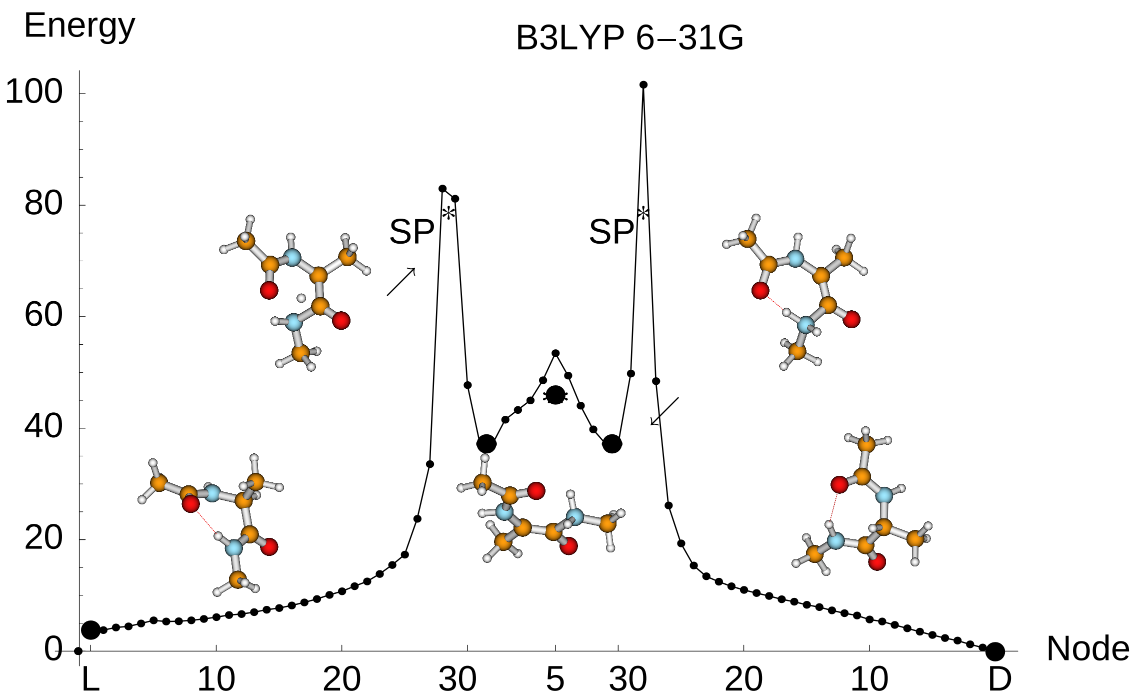


Figure: Approximation of the energy profile in kcal/mol of a Newton trajectory by the GS method for alanine dipeptide between the minima of the L-enantiomer of $C_{7\,ax}$ (left) and the corresponding D-form (right). Bullets are minima, stars are saddle points.

Introduction: Search of a TS by a 1D curve, by the "reaction path"

Finding a transition state (TS) between two minima is one of the most important problems in chemical physics, since it provides reaction rate constants in combination with the transition state theory. Guesses of a TS and its support, the minimum energy path (MEP) are sometimes very difficult in a multi-dimensional potential energy surface (PES), and poor guesses often fail in reaching a special TS. Thus, it is very important to establish a technique which may lead to an estimate of a TS region without an initial guess of the TS. The growing string method (GS) [1] has recently been developed for this purpose. A long list of methods shows that finding a TS is difficult because of its mathematical condition on the PES. It is a stationary point, a saddle point, with a single negative Hessian eigenvalue. For finding a TS on a PES, the energy should be maximised in one direction and minimised in all other directions. A stable equilibrium (EQ) structure can be rather easily located by simply minimising the potential energy. In a high-dimensional space, the maximising direction for the saddle point is difficult to determine. Various techniques have been developed for finding a TS, see the references in [2, 3].

Since the structures of the reactant (R) and the product (P) are known, a test chain can be used between them. On the other hand, if either R or P is unknown, one-dimensional (1D) search curves starting from a known structure, and walking locally on the PES under a clever directive, can be used for finding a pathway to the other unknown structure via a TS. Various test chain methods have been used, such as the nudged elastic band method [1, 4, 5], string methods [6, 7, 8, 9], and others, see [2]. In the conventional TS search methods, an eigenvector direction of the Hessian matrix is often assumed for uphill walking toward the TS. The eigenvector following method [10] in quantum chemical programs (cf.[11]) can be used to locate a TS if it exists around an eigenvector direction initially chosen. (A 1D search curve is employed here more implicitly; it has a loose connection to the gradient extremal [12].) Other 1D search curves used in the past are Newton trajectories [13, 14], or even the gradient extremals [12, 15].

Our growing string method (GS) [8, 9] does not search a steepest descent curve, the intrinsic reaction coordinate (IRC), however, it employs any Newton trajectory (NT). It is a curve where the gradient (G) of the PES is always a pointer to a fixed direction [13] called the search direction, $\mathbf{r}$. The NT has roots in the old known method of a driving coordinate, however, it avoids the breakdown of that method: to be not always continuous, to jump uncontrollably across the PES behind a turning point. There is a recent paper suffering under such problems [16]. The breakdown of the old method is circumvented by another mathematical definition of the NT which we will follow. Note that the NT and the driving coordinate fall together in the cases where the old method works well. They are not equivalent behind turning points. The NT is fitted up with a solid mathematical foundation [13, 14]. The property of an NT (to have a fixed gradient direction) is realisable by a projection of the gradient. We search a curve $\mathbf{x}(t)$ where t is the parameter, and it has to hold for all t from reactant over the TS to product minimum

$$\mathcal{P}_r\, G(\mathbf{x}(t)) = \mathbf{0}\ . \tag{1}$$

The projector $\mathcal{P}_r$ can be a constant matrix of rank n-1 if n is the dimension of the PES. We use a column vector $\mathbf{r}$ for the projection. It has to be a unit vector. The projector which projects orthogonal to $\mathbf{r}$ is

$$\mathcal{P}_r = \mathbf{I} - \mathbf{r}\cdot\mathbf{r}^T\ . \tag{2}$$

$\mathbf{I}$ is the unit matrix and $\mathbf{r}\cdot\mathbf{r}^T$ is a dyadic product. $\mathbf{r}$ may be fixed for a Newton trajectory adapted to any given search problem. Every regular Newton trajectory connects EQs with an index difference of one [17], thus, for example, Newton trajectories connect minima and TSs [18].

This paper is an application of the GS method of the author [8, 9] to a high-dimensional PES with 486 dimensions. Before we apply the method to a "toy"-problem, to alanine dipeptide. The results reported for the polyalanine are not fully satisfactorily; nevertheless, they show the way to a possible application of the method. At the state of affair, no comparison is given to other methods, cf. [1, 4, 5, 19]. This should be done in a following work.

Isomerisation reaction of alanine dipeptide

Alanine is one of the most common proteinogenic amino acids. Applications of the GS method are done to alanine dipeptide (Acetyl-L-Ala-NHMe),

$$\mathrm{CH_3CO\text{-}NHCHCH_3CO\text{-}NHCH_3}. \quad \mathbf{1}$$

We use here the usual nickname "dipeptide", though it is a single amino acid with one Cα-atom only in the centre, and two caps. But it contains two peptide bonds, C-N, between the two peptide units, the row of atoms O,C,N, and H. Work [9] uses NTs to represent RPs for the 66-dimensional configuration space of an alanine dipeptide isomerisation, at the lower 3-21G level. The quantum chemical level of DFT calculation, B3LYP/6-31G, is used here to compare the effort with other results. (To higher level calculations of EQs of alanine dipeptide see ref.[20].) The level B3LYP/6-31G gives a simpler PES than in the 3-21G case [9] because an intermediate minimum and a corresponding SP disappear here. Maeda and Ohno [2] note that a TS between the two conformers C_5 and C_{7ax} of alanine dipeptide is deviated from the midpoint between them. The choice of $R = C_{7ax}$ and $P = C_5$ was made to obtain higher computational efficiency according to the Hammond rule. The scaled hypersphere search method exactly finds an early TS [2]. Applications of another method, the growing string (GS) with a nudged elastic band were also recently done to this system [1]. The method is a MEP optimisation. It searches the steepest descent curve (IRC) from the saddle (without guessing the SP in before).

The proposed GS-NT method [8, 9] uses the Newton trajectory (NT) of eq.(1) to any start direction (for example, $\mathbf{r}$ can be the direction between the given two minima). The NT is the leading curve approximated by nodes. After a predictor step the GS-NT method employs Nocedal's L-BFGS method [21] in a projection version [9] to find a zero of eq.(1) orthogonal to the search direction, $\mathbf{r}$. For every molecule emerges the task to determine a useful convergence condition, ϵ, for the norm of the projected gradient (1) of a current curve point. (Or for every single component of the projected gradient.) ϵ should be strong enough to stop the corrector not too early before the RP node. If it is too strong, we waste a lot of computer time. Then the corrector does ineffective calculations, so-called zigzagging, under a small sliding down along the direction of the gradient.

For alanine dipeptide, **1**, we had found that $0.002\,5$Å$\leq \epsilon \leq 0.007\,5$Å are possible values for the threshold [9]. Here, we use ϵ=$0.003\,5$Å for three calculations with different numbers of string nodes: 5, 7, and 9. We use the linear combination (of curvilinear z-matrix coordinates in this example, later we also use Cartesians directly) of the current chain point and the final point [9]. We search an RP which should connect the initial minimum, $\mathbf{x}_{ini}$=$\mathbf{x}_0$, with the end $\mathbf{x}_{fin}$=$\mathbf{x}_{m+1}$ by a chain of m nodes $\mathbf{x}_k$. We calculate successive nodes beginning at the initial minimum by a number of copies of the original molecule, somehow interpolated between the current structure, $\mathbf{x}_k$, and the product structure. Since $\mathbf{x}_k$ is an approximate node on the RP obtained by the last corrector loop, we choose a next guess point, $\mathbf{y}_{k+1}$, of the string between the actual node and the final minimum by the predictor step

$$\mathbf{y}_{k+1} = \lambda_k\,\mathbf{x}_k + (1-\lambda_k)\,\mathbf{x}_{fin}\ , \quad \lambda_k = \frac{m-k}{m+1-k}\ , \quad k=0,..,m-1\ . \tag{3}$$

When the RP was a straight line, the linear combination would give equidistant nodes on the line if we set back successively $\mathbf{x}_{k+1} = \mathbf{y}_{k+1}$ without the corrector and repeat ansatz (3). However, coordinates are often z-matrix coordinates, and the RP is usually curvilinear, so the predictor nodes are not necessarily fully equidistant. Indeed for long pathways with strong curvilinear RP, to avoid excessive rugged regions of a PES, often we do not use the linear combination for Cartesian coordinates of the two structures of the molecule. We use devise (3) for internal coordinates in the predictor. Of course, the condition is: the z-matrix has to be the same in $\mathbf{x}_k$ and $\mathbf{x}_{fin}$. But the corrector uses Cartesians. We do a conversion of all the distances and angles into full Cartesian coordinates. After setting the zero of the coordinates to the centre of mass, the set is then given into the Gaussian03 (over a GS-shell [9]) which performs energy and gradient calculation in all degrees of freedom. Then the GS corrector loop using the CG+ method is done in Cartesian coordinates. Each resulting structure individual $\mathbf{x}_k$ (of the corrector) is again converted back from the Cartesians into z-matrix coordinates for the next predictor, and the next iteration commences.

Figure 1 is a result of the GS method with 9 nodes. We need the following number of gradient calculations:
(i) For 5 nodes, we need 212 calculations of gradients for the RP approximation with ϵ=0.003 5Å , and starting at the highest point, here node 3, we need 40 steps in a Berny optimisation of the exact SP. A TS was deemed to have converged when the following values fell below the given convergence condition of the Gaussian03: Maximum Force 0.000 45, RMS Force 0.000 3, Maximum Displacement 0.001 8, and RMS Displacement 0.001 2. All values are in atomic units.
(ii) For 7 nodes, we need 279 gradients for the RP approximation, and from the highest point, node 4, we need 49 gradients in a Berny optimisation of the exact SP.
(iii) For 9 nodes, we need 277 gradients for the RP approximation, and from the highest point, node 5, we need 30 gradients in a Berny optimisation of the exact SP.
Figure 2 additionally shows the protocol of predictor- and corrector steps along an energy profile. Already the small bend per predictor of alanine dipeptide, **1**, in z-matrix angles by eq.(3) under 9 nodes causes a large distortion of the energy of the current point after the predictor step. (The effect would be much larger for a straight initial chain between $C_{7\,ax}$ and C5.) It can be seen how the corrector steps in one loop lead back to the profile over the NT.

The Berny procedure approximates the eigenvalues of the Hessian matrix of the current point by gradients of preceding points. It is worthy of a remark that the Berny procedure reports on its way down throughout a very small number of negative eigenvalues (one to three, mainly 2), indicating indeed a nearby valley ground. It is in contrast to a different experience in ref.[1].

Pathway between two quasi-enantiomers of $C_{7\,ax}$ of alanine dipeptide

A next application of the GS method is done using the $C_{7\,ax}$ structure of alanine dipeptide too, see Fig. 3. However, we try to determine a pathway through the high energy mountains of the system, going besides the low valleys of conventional isomerisation calculations. We choose the enantiomerisation from L-form to D-form, or vice versa. The molecular structures are given by inlays in Fig.4. The energy of the minimum at B3LYP 6-31G level of the usual L-enantiomer is -495.705 91 au, and that of the D-enantiomer is -495.709 80 au. Thus, the D-form is 2.44 kcal/mol more stable than the L-form by steric asymmetry, already in the alanine dipeptide, **1**. It holds in contrast to the simple alanine where the two enantiomers are quite equal (c.f. the discussion of the "equalness" and references in [22]). The definition of mirror symmetry, of real enantiomers is not fully given for the alanine dipeptide, that the two mirror forms are identical under non-chiral conditions. We have only two quasi-enantiomers.

The enantiomers concern the Cα atom C8 in the 7-ring, (with atom number 8 in a consecutive numbering if starting at the Acetyl cap, see the z-matrix [23] and Fig. 3) and its bond to the H atom: C8-H11, as well as the bond to the CH_3 group: C8-C13. The two groups lie alternately outside the plane of the 7-ring. In the naturally treated L-form, the CH_3 group stands skewly over the 7-ring, where the H11 stands off from the ring but lies nearly in the plane of the ring. In the quasi-D-form, contrary, the C8-H11 bond stands over the ring where the Me group stands off from the ring. A simple turn of the positions of the two bonds would have to pass nearby the bonds of the ring neighbours of C8, the C8-N7 and C8-C9 bonds. However this is sterically forbidden. Enantiomerisation is a first-order reaction. It means interconversion of one enantiomer into the other. To propose a possibility to enantiomerisation of **1**, we follow an idea of Bai, Wang, and Min [24] that the Cα-H bond would break, thus, the H11 has to move away from the Cα. The remaining CH_3 group can turn into any position and the H11 (or another H atom) may come back to complete the other quasi-enantiomer.
We have tested the neighbour nitrogen N7 of C8, or the other nitrogen N17 of the ring, the next neighbour C9, the next O10, as well as the counterpart oxygen O4 of the ring, to play the role of the parking place of Hα. Atom C9 can not hold the H11, it immediately goes back to the Cα under an optimisation.
(i) Perhaps the most intriguing example is provided by the N atom of the cap, N17, being an atom of the 7-ring. The kind of proceeding is the following: we cut the H11 from C8=Cα and put it to the atom in question, N17. A search for a minimum with Gaussian03 is easily to do: it gives the minimum at -495.649 98 au. Then we try to find a GS pathway in Cartesian coordinates between the two minima using the directional vector between the two structures for the search direction, $\mathbf{r}$. Along the RP nodes, the $\mathbf{r}$ is additionally adapted in direction to the final structure [9]. The threshold chosen for the corrector was 0.005Å, and, because we need a big rearrangement of the backbone, a big number of 30 nodes is used. The energy profile of the pathway found is shown in Fig.4. The GS method first rearranges the backbone of the molecule. At the beginning, the one H atom which should do the jump, is confronted with 21 other atoms of the molecule which also move something. Thus, the directional part of the H11 in the search vector, $\mathbf{r}$, is small. But the covalent bond of H11 to Cα is strong. The GS is going slightly uphill in energy along 24 steps, up to reaching the region with a stronger slope where the jump of the H atom has to be done. There the threshold ϵ is made weaker: it is multiplied per step by the factor of 1.1. The TS itself is calculated by the Berny method of Gaussian03 [11] using again the node with the highest energy of the string. The SP is at energy

of -495.583 57 au. In inverse direction, an analogue pathway has gone uphill beginning with the quasi-D-form. The TS here has the same energy, as well as the enantimeric minimum of H11 at N17, which has the same energy as the minimum in L-form.

There are at least three relevant possibilities for the former H11 to come to anchor at N17: it may stay out of the plane of the 7-ring where the former H-bond with H20 continues to exist, that is a mimimum. The 7-ring may flatten out totally, its H-bond is solved, and the two H atoms may lie symmetrically, below and over, to that plane: that intermediate is an equatorial structure between the two minima. It has a symmetric shape: all heavy atoms and four H-atoms are in the plane of symmetry, σ_v, whereas four pairs of H-atoms can be reflected through the plane of symmetry. Or, the former H11 may become the H-bond of the ring, and the other H atom, the H20, may go outside the ring: that is the other minimum being the real enantiomeric to the former. The pathway between these two minima is also approximated by an NT in the centre of Fig.4 by 8 nodes. Thus, the full pathway of enantiomerisation could develop, in this kind: First, after a long, often contrary arrangement of the backbone, the H11 from Cα jumps to N17, then it exchanges its order there with the other H atom by a further rearrangement of the backbone which is going through the flat σ_v-symmetry, and becomes the H atom of the H-bond. And last, the remaining H20 atom may jump back to the Cα at the other side of the 7-ring, and the backbone also rearranges to the structure of the D-quasi-enantiomer.

The strings of the example cover long pathways over the mountains of the PES of alanine dipeptide. Because the PES is known to be rugged, we had chosen a large number of nodes. However, the pathways emerge to be direct ascents from the $C_{7\,\mathrm{ax}}$ minima to the corresponding TS of the H11-jump. Maybe the NTs are not always valley pathways, but they go along the slope of the corresponding hills. The computational effort of the strings in Fig.4 is: 256 corrector iterations for the first part, where, of course, more iterations are done for the steep ascent to the transition state region; then 94 iterations for the second part between the two enantiomeric minima, and 340 iterations for the last part where again more iterations are done along the steep ascent. In average we have less than 10 iterations per corrector loop.

(ii) The next case treated is the transition of the H11 atom to the next neighbour but one, O10 of C8. There is a minimum at -495.661 46 au. The energy profile over two NTs is given in Fig. 5. The calculation of the NTs is again done with 30 nodes, and the threshold for the corrector is 0.009 Å. Here, the TSs calculated from the two highest points of the two GSs are either a place where H11 is already at O10 at an energy of -495.603 08 au, or it is the place of H11 at the atom in between, the C9 atom at -495.564 30 au. Both pathways are not symmetric. The backbone has to be arranged in different kinds for the two pathways. The number of corrector steps was 256 for the left part starting from the L-enantiomer, and 147 for the right part starting from the D-enantiomer.

(iii) The case H11 at N7 is a different story. There is a minimum at -495.598 48 au and a similar TS at -495.598 47 au (with some alternate backbone angles). However, there are further TSs, and at least, the structures tend to become unstable: the covalent bond between N7 and C3 can be broken, or the Me group can additionally leave the Cα, or further cases can happen. Results of corresponding calculations should be given in a next paper.

The cases (i) and (ii) result in two or three TSs for the enantiomerisation in question, by explicit pathways of 30 nodes. We did not search further possible, lower TSs or intermediate minima. The existence of a 30-nodes pathway demonstrates in every example that the given TS connects the corresponding minima, thus is indeed a TS. The used search direction of the NTs leads to the success, however, we cannot exclude further parallel pathways over different TSs.

Bending back of a bent chain of 15 alanines

How a 1D polypeptide chain folds into a 3D biologically active form is one of the greatest challenges in life science. Several decades ago Levinthal [25] estimated that during the actual folding process of a protein, it explores only a tiny fraction of the conformational space. The latter grows exponentially with chain length. However, proteins in real life fold into one final form. This apparent discrepancy can be resolved by the assumption that (i) folding of a real-life proteine is not a random walk on a flat PES, instead (ii) it proceeds from unfolded states via intermediate and TSs and is then funneled towards the final state [26, 27].

Linear and bent helix

As a substitute for a proper benchmark, we investigate a polyalanine, N-acetyl-[L-alanil]$_{15}$-N′-methylamide,

$$CH_3CO\text{-}[NHCHCH_3CO]_{15}\text{-}NHCH_3. \qquad \mathbf{2}$$

The 15-mer sequence is abbreviated by Ac(Ala)$_{15}$NHMe. Ac is the acetyl group, -$COCH_3$, protecting the N-terminal, and Me is methyl, -CH_3, where the group N-methyl amide, -NHMe, protects the C-terminal. The chain is capped, because otherwise charges at the termini reduce the helix content [28]. The chain has 162 atoms, the PES extends over 486 dimensions. In polyalanine the main chain N-Cα and Cα-C′ bonds are relatively free to rotate. These rotations are represented by the torsion angles ϕ and ψ, respectively, called Ramachandran coordinates [29]. A helix combines a linear translation with an orthogonal circular rotation of the chain parts [30]. But polyalanine helices are additionally stabilised by periodic H-bonds. The helix of alanine parts can be described by the number of residues per helical turn and the number of atoms contained in the corresponding ring made by the H-bond. This study largely confirms the notion that the secondary structure of polyalanine is mainly an α-helix former in vacuo, signifying a 4_{13}-structure (the H-bond reaches from residue i to $i+4$, and a 13-membered ring of atoms is formed), at least at the B3LYP/6-31G level: The fully helical state forming a minimum that we were able to optimise contains eleven H-bonds of a right-handed α-helix character, and three H-bonds of a 3_{10}-helix, see Fig. 6 and Tables I and II. We use the H-bond indicator of the program MOLDEN4.4 [31]. Thirteen H-bonds could be counted in a theoretically ideal model in a consecutive 4_{13}-structure [32]: the CO$\cdots$HN bonds span four alanine parts. The caps provide the first and the last H-bonding sides. With the C=O of the acetyl cap at the N-terminus plus the 15 alanines plus the N-H at the cap of the C-terminus, there are 13 possibilities: three free N-H remain without an H-bond to an oxygen at the beginning, and three free C=O remain without an H-bond to nitrogen at the lower part. However, minimisation of the structure disturbs the total α-helix symmetry and enforces another H-bond diversification. At the C-terminus the last two possible 4_{13}-bonds are replaced by two 3_{10}-bonds. This is the usual case: the 3_{10}-helix part is always short and frequently occurs at the termini of regular α-helices. The intermediate oxygen-atom, O61, (with numbering first of the backbone beginning at the N-terminus) of the fifth alanine from below (residue 11) is free, the CO group is buried in the interior, and it separates the two parts of the helix: a long α-helix and two residues of a remaining 3_{10}-helix, see Fig. 6. So, the two helicity archetypes are both represented. It confirms CHARMM calculations where the C-terminus also prefers 3_{10}-helical structures [33]. Experimental results [34, 35] also suggest that alanine-peptides fold as a mixture of 3_{10}- and α-helix. The buried CO group contradicts the requirement that in the final state all helical H-bonds have to be satisfied [33]. Thus, this artificial model helix, **2**, is frustrated. There is not a single conformation of the chain which optimises all the interactions at the same time [36]. The N-terminus additionally contains one H-bond reminiscent of a 3_{10}-helix that links the terminal acetyl C=O to the N-H from the third alanine from above, see Fig. 6, compare an analogous result of a DFT/AM1 ONIOM calculation [37]. As the same C=O also interacts with residue 4 to form a normal α-helix H-bond, the oxygen atom forms a pair of H-bonds. The feature makes a bond bifurcation. (The hybrid $\alpha/3_{10}$ helical structure disappears at the lower B3LYP/3-21G level. This seems also to be the case in an ECEPP/2 potential [38].)

We use the new found minimum of a bent state,

$$\text{Ac-(Ala)}_7\text{-Ala-(Ala)}_7\text{-NHMe}, \qquad \mathbf{3}$$

see Fig. 6 and Table I, to answer the question: Can we determine the transition path from the bent structure back to the erected helical minimum? The bent state here is used as reactant, and the fully helical state is taken as the product configuration. We assume that the bent-to-straight helix transformation of Ac(Ala)$_{15}$NHMe can be described by an overall two-state model, and we also assume that the two minima are representants of deep funnels. The pathway is assumed to be a long pass. However, the floor of the pass may be rugged by a lot of small wells given by changing H-bonds, or turning CH_3 groups in any side-chain-dimension, for example. But for a first extrapolation, we assume that the landscape is smooth, with a quite large energy gap separating the (native) straight helix from the bent conformer, cf. [39], and that the overall slope of the PES is such that the molecule is easily driven down to its sink states, i.e. the PES should be funnel shaped, cf. [40]. The assumptions will be sanctioned by the result of our calculations. For a first view in z-matrix coordinates, we treat an initial chain of test nodes with ansatz eq.(3) but without corrector steps, c.f. ref.[8]. The energy profile of these chain nodes rises up to 60 kcal/mol over the initial bend.

The bent state is given by an individual turn formed over residue 8, being the hinge, thus by a large change of the torsional angles around the Cα-atom in residue 8. A milder distortion concerns some neighbouring torsional angles, see Table I. Full geometry data are available, see [41] and supporting info. Because the Ramachandran angle ϕ_8 is strongly changed, the bent helix is a coil in the usual understanding [42]. (But it is not a random coil; so we name it a bend.) The other changes in the remaining ϕ_i ($i \neq 8$) and ψ_j are moderate and they usually change alternately: if ϕ_i becomes larger, the corresponding ψ_i becomes less. For example, it is $dif(\phi_{10}) = -11.65^o$, but $dif(\psi_{10}) = 28.54^o$. Thus, the two parts of the bent helix will continue to emerge like straight helices. A speciality of the minimum structure of the bend is the large torsional turn of the side chains, the Cα-Hα pair in residue 8 (dih120) by 130.4^o, and an analogous turn of the Me group by 128^o (dih40).

The minimum of the bent helix that we were able to optimise contains twelve H-bonds forming two helical parts. However, it again contains one H-bond bifurcation with a second bond of 3_{10}-character that now links the oxygen atom, O46, in residue 8 additional to the α-bond, but after the bend of the structure in the C-terminus part. The end of the C-terminus part is a 3_{10}-helix with three H-bonds, where two of the 3_{10}-bonds of the linear helix have survived, and a new contiguous 3_{10}-bond is formed additionally. Thus, only two 4_{13}-bonds remain in the

C-terminus part. In the N-terminus part remains the 3_{10}-bond at the beginning, but the bond bifurcation is dissolved because the corresponding 4_{13}-bond disappears. (It has the length of 3.38 Å .) Two old 4_{13}-bonds follow, however, after that the character again changes. The oxygen in residue 5 is free and its CO group is buried, and the next two H-bonds before the hinge are 3_{10}-bonds. At least, at the hinge itself, there emerges an extra $(i, i+2)$-H-bond in a seven-membered ring of atoms, compare [43, 44]. It is formed by a 2_7-ribbon in boat shape like the minimum $C_{7\,ax}$ of alanine dipeptide, see Figs. 1, 4. At all, there are five buried groups which may be responsible for the ≈ 12.74 kcal/mol higher energy of the bent helix in comparison to the straight helix (besides the steric strain of the bend). These groups are CO in residues 3 and 6, NH in residue 4 in the part of the N-terminus, and one CO and one NH in residue 10 in the part of the C-terminus. Podtelezhnikov and Wild reported [45] the first computational observation of a stable kink between two 3_{10}-helices accompanied by an H-bond bifurcation. Here, the hinge of the bend is also between 3_{10}-bonds. But it emerges additionally a 2_7-ribbon at the hinge.

Eight H-bonds are nearly uniform in both of the structures of interest, but three have to change from 4_{13}-bonds into 3_{10}-bonds, three 4_{13}-bonds disappear, one bifurcation of a 4_{13}-bond into a hybrid $\alpha/3_{10}$-bond is to be built newly, and the extra H-bond in a seven-ring has also to be created. All the changes of the bent structure have to be retracted along the searched reaction coordinate. The test case was selected, because polyalanines with $n < 15$ tend to build an α-helix formation in the vacuum case, see for $n = 10$ refs. [38, 46]. Up to $n = 10$ the energy of the α-helix formation increases with n in the short range, as well as in the long range, electrostatical interactions [47]. Polyalanines with $n \geq 20$ can show bundles of two short helices, linked by a turn section. They are more stable than the single straight α-helix [48]. We assume that a helix with $n = 15$ alanines meets the intermediate, cf. [37, 49]. A further hint is that the existence of metastable states was obtained in molecular dynamical calculations for a chain of 15 alanines [50].

Polyalanines are an extensively studied sort of molecules which show a funnel structure of their PES. The discussion began with the Ac(Ala)$_3$NHme [51], and already in Ac(Ala)$_8$NHme [52] a great system of funnels was found. The treatment was continued to Ac(Ala)$_{12}$NHme and Ac(Ala)$_{16}$NHme [49] where again diverse funnels exist. (These reports work with empirical potentials.) We can hope that funnels determine also the PES for the Ac(Ala)$_{15}$NHme.

A remark on dimensional restrictions

There are many reasons for the use of a minimal number of dimensions of the PES [53]. However, in the case of Ac(Ala)$_{15}$NHme, the use of restricted dimensionality framework only, which may be suggested also by some lines of Table I, does not give the resulting MEP (see profile in Fig.7). The reader may find a similar discussion and quantitative results in ref.[54]. Only the treatment of a wide dimensionality gives the possibility to find the low transition region of this example. Along the wandering of the different corrector loops (a 2D projection is Fig. 8) there are a lot of changes in different regions of the molecule: the folding requires the formation of a lot of noncovalent interactions and the synchronisation of a large number of degrees of freedom, see below. Restricted dimensionality would exclude such possibilities from the beginning. One may hope to go along the same global valley, but only a little higher in energy, with a resticted dimension. However, this must not be the case: the constraints may also cause a deviation of the program into other valleys of the PES and a quite different course, at all. Of course, already the concept of the RP may be understood as a simplified description of reduced dimensionality.

But we use the RP as a 1D leading curve in the full dimension of the problem.

Calculation of the reaction coordinate

We employ an *ab-initio* strategy to achieve the RP in the form of an NT. The quantum mechanical calculation is tractable, for this model protein structure, at the current state of computational possibilities. We set up the GS machinery with a chain of 10 nodes to connect the two structures of interest. For the predictor we use the linear combination, devise (3) for internal, z-matrix coordinates. The book-keeping of the z-matrix is a formidable task, especially it is so for the torsional angles. Each resulting structure individual of the corrector in Cartesian coordinates is converted into z-matrix coordinates maintaining the numbering. Non-unique torsional angles are automatically adapted to the final structure, by addition or subtraction of 360^o, correspondingly. One can compare Table I for some important cases. Employing the z-matrix, the predictor (3) is done, using the z-matrix of the final structure, the straight helix. The predicted point is fed into MOLDEN4.4 [31] to inspect it in every case, and then it is converted back from the internals into Cartesian coordinates. The set is moved into the centre of mass, and it is given into the Gaussian03 (over a GS-shell [9, 41]) which performs energy and gradient calculation in all degrees of freedom. Energy output and gradient output of the Gaussian03 in Cartesian coordinates are used in the CG+ program of the GS code [9, 21].

Note that the large turn from bent helix to straight helix of Ac(Ala)$_{15}$NHMe, treated in Cartesian coordinates by ansatz (3), would give a forbidden clash: the mutual intersection of parts of the molecule with impossible high energies. We found that the use of internal coordinates for the bending back to the final straight α-helix is forcing. In the z-matrix coordinates, the initial chain of nodes with ansatz eq.(3) is clash-free. It is remarkable that the energy of the chain is at node 4 about 60 kcal/mol over the bend, or 73 kcal/mol over the α-helix. We did not try to use this test chain. Along the GS chain, quite better, the energy of the current structure changes maximally with the predictor-jump up to 10 kcal/mol due to the linear combination of eq.(3). Of course, the energy then falls again back under the corrector.

The maximal step length of the corrector steps of the employed CG+ method [21] is artificially restricted to 5.0 Å in Cartesian coordinates, which is large in comparison to the dipeptide example above. But steps here have to spread over more coordinates in such a large molecule. Usually, the corrector steps are shorter. The length is calculated by the line search device of the CG+ method [21]. The result of the GS search is shown in Fig.7. A chain of ten nodes of the growing string is calculated between the minima, where the corrector did need the following number of iterations: 10, 32, 20, 45, 17, 26, 28, 19, 6, and 6. The threshold for the corrector convergence was at the beginning 0.002 5 Å up to the putative TS before node 6, but on the decreasing part of the pathway after node 6 it was lowered to 0.003 5 Å to economise the calculation. Corrector loops stop here when all components of the projected gradient are smaller than the threshold. Note that the diverse corrector steps may go into different directions on the high-dimensional PES which are to add, or to subtract to the corresponding predictor step. For a projection into the two Ramachandran angles ϕ_8 and ψ_8, the action of the corrector is shown in Fig. 8. The thin line there is the approximated pathway of the bending, and dots are the steps of corrector loops after the corresponding predictor jump from the given thin line. The given numbers are the nodes. Node 7 is placed downhill of the strong slope behind the border of the funnel of the linear helix. There the corrector wanders a large distance before convergence. Thus, the exact scaling between the nodes 1 to 10 is not explicitely given. The nodes in Fig.7 are only current numbers.

The search for a TS

Calculated in the full dimension, the corrector loops still need a lot of small steps for convergence to the given profile in Fig.7. A low lying path could be detected, if we accept the nodes of the corrector, with an energy of the highest node of 12.5 kcal/mol over the bent structure and 24 kcal/mol over the straight helix. For reference, $k_B T$ at 300 K is ≈0.6 kcal/mol, so it is a significant barrier when there is also a TS nearby. The profile looks like a broad, moving transition state [55]. However, we could not find any detectable intermediate minimum. Treatments of Berny optimisations in Gaussian03 for the search of an intermediate minimum were not successful if one starts at nodes 3 to 6. (We used small steps: we set the parameter MaxStep=3. Thus, there are no jumps of the course downhill.) The optimisations slid down into the left (nodes 3 and 4) or right (node 5 or 6) funnel. There does not seem to be any place for a local trapping at the top of the barrier. Only if starting at node 2 on the initial slope of the RP, we find a local minimum of a bent structure without the 2_7-ribbon at the hinge, cf. ref. [45], at an energy of -3957.337 48 au, thus, 1.57 kcal/mol over the bend with the developed 7-ring.

Searches for TSs, where we make any effort to find one, have an analogous outcome. Berny optimisations in Gaussian03 for a TS do not find a structure existing nearby. Again we start at nodes 3 to 6, and we use small steps. All optimisations slid down into the left (nodes 3 or 4) or right (node 5 or 6) deep sink to SPs of a higher index, mainly index 2. We can guess that the TS should be between nodes 4 and 5, but we could not pick up a proof of the claim.

It is clear that bending along the angle ϕ_8 must be a part of every reaction coordinate which describes the bend to α-helix transformation. For small molecules it is a fixed assumption that simple reactions have a clear pathway on the PES. In contrast, proteins have many degrees of freedom. Many workers believe that conformational changes of proteins may be imperfectly captured in 1D or 2D reaction coordinates [56]. In the theory of peptides, the next stage of a definition is going from TS to a transition state ensemble of points, or a transition tube, which are given by the probability of ≈0.5 to fall either into the folded funnel, or into the funnel of the bend [56, 57]. Using this definition, we can decide that node 4 may still belong to the sink of the bend, but node 5 may already belong to the sink of the linear helix. In an analogous test, we examine points $\mathbf{y}$ of a linear combination between nodes 4 and 5

$$\mathbf{y}_\lambda = \lambda * node\mathbf{4} + (1-\lambda) * node\mathbf{5} \ . \tag{4}$$

Eight consecutive test runs of a TS optimisation with MaxStep=5 are calculated with values $\lambda=\frac{1}{2}$, $\frac{3}{5}$, $\frac{73}{120}$, $\frac{37}{60}$, $\frac{38}{60}$, $\frac{39}{60}$, $\frac{2}{3}$, and $\frac{4}{5}$, see Fig. 9. The cases $\lambda=\frac{1}{2}$ and $\lambda=\frac{3}{5}$ fall directly down to the sink of the erected helix with negative $\phi 8$ values. The case $\lambda=\frac{4}{5}$ falls down a longer distance of energy inside a small region of $\phi 8$, but then it escapes into the sink of the bend. Tests are erratic when they are inside of that interval. The TS search goes monotonously downhill in energy but meanders in $\phi 8$. A TS is not met in this region. The one curve with points in Fig. 9 is that of $\lambda=\frac{39}{60}$, here the first 160 points are shown. It escapes to the bend under larger negative values of ψ_8 and ϕ_9. It crosses from right to left the curves of $\lambda=\frac{38}{60}$ and $\lambda=\frac{37}{60}$ which fall into the bends sink, as well as the curve of $\lambda=\frac{73}{120}$ which falls into the sink of the α-helix. Thus, doing an interval inclusion does not bring up the solution: the points near a putative TS-value of $\phi 8$ at ≈16^o show at the beginning a more or less clear trend downhill in energy along a more or less equal value of the leading coordinate, $\phi 8$. Then in hopeful cases a meandering inserts, and after that, under further, mild downhill steps of energy, the optimisation goes into only one direction of $\phi 8$. Then it seems that we already are in one of the two funnels. However, the

Berny procedure still indicates the existence of between five and eleven negative eigenvalues. Using the illustration of the behaviour in Fig. 9, we may pessimistically guess the value of the barrier of the RP ≈5 kcal/mol below node 4, because down to that level the optimisation goes in larger steps, but after that level the energy steps are smaller. It seems that the optimisation has reached the floor of a high-dimensional valley ground, and, at the same time, a corresponding broad ridge in four to ten dimensions. The direction of the one shown coordinate $\phi 8$ becomes erratic. Its direction is overlapped by other, stronger dimensions. Optimisation steps seem to do some zigzagging which is known to take place on the floor of a valley. Of course, here the zigzagging includes other degrees of freedom too. Also the norm of the gradient of the PES becomes quite small at the given level.

The fiasco of some searches for the RP barrier suggests, on the other hand, that any pathway to the two global minima leads along a funnel-like energy profile. The set of configurations of the given approximation of the RP seems to be free of intermediates. In the isomerisation reaction of this polyalanine we did not find any hint to some adjacent TSs and adjacent intermediates for the forming or opening of single H-bonds, in contrast to ref. [58], Fig. 1 there, and ref. [27]. Though we found two quasi-peaks in Fig. 7, nodes 3 and 6, we have to assume that the underlying energy profile should have a single highest barrier along the progress coordinate, between nodes 4 and 5, with the value of $\phi 8 \approx 16^o$. However, our RP is not the floor of a valley. It is a mixture of valley and ridge character, the latter holds on for ten directions at the top.
We may speculate that more nodes and a stricter threshold for the convergence of the corrector will lead to a result for the barrier.

Discussion of dissolving or forming of H-bonds along RP

The forming of an H-bond in polyalanine is a long studied problem [59]. The energy of the donor-acceptor interaction of an isolated bond, NH···O=C, is estimated to 5-6 kcal/mol in a 13 mer α-helix [60], to ≈ 6.2 kcal/mol for α-bonds in Ac(Ala)$_n$NH$_2$, with $n = 2$ to 18, and ≈ 7.5 kcal/mol for 3_{10}-bonds [37], and to ≈ 5 kcal/mol in Ac(Ala)$_{15}$NHMe for the bond at the N-terminus [32]. A different result is an average energy of 2.13 kcal/mol for Ac(Ala)$_{10}$NHMe for α-bonds [61]. The 7-ring in Ac(Ala)$_2$NHMe is estimated to 5.73 kcal/mol [46].

The reaction path for Ac(Ala)$_{15}$NHMe which we were able to approximate does not give the explicit possibility to estimate the energy of single H-bonds. Other non-covalent forces are also comparable and may superimpose the H-bond changes. The determination of the strength of a specific atomic-atomic interaction is often quite difficult, as it may simply be impossible to separate the two atoms to large distances without breaking other bonds [62]. However some hints can be taken, see Table II. First of all, however, we have to observe that there is no sudden forming or dissolving of an H-bond. For example, the jump from 3_{10}- to α-bond in residue 14 in the C-terminus goes along the following bond lengths: the 3_{10}-bond from O61 to H149 opens like
$(2.02 \rightarrow 2.05 \rightarrow 2.08 \rightarrow 2.12 \rightarrow 2.13 \rightarrow \underline{2.19} \rightarrow 2.25 \rightarrow 2.32 \rightarrow 2.4 \rightarrow 2.5 \rightarrow 2.58 \rightarrow 2.68)$Å
where the 4_{13}-bond from O56 to H149 rises up with
$(3.97 \rightarrow 3.8 \rightarrow 3.66 \rightarrow 3.5 \rightarrow 3.38 \rightarrow 3.21 \rightarrow 3.02 \rightarrow 2.82 \rightarrow 2.66 \rightarrow 2.51 \rightarrow 2.37 \rightarrow \underline{2.24})$Å.
The underlined values are distances where MOLDEN4.4 reports the dissolvation or the new existence of the corresponding H-bond (defined from 1.5 Å ≤ distance O···H ≤ 3.15 Å and bend angle O···H-N falling in the range of 35^o from linearity). The line of decision is more or less a question of a formal definition, compare [32, 50, 59, 61, 63, 64]. Only small changes of the Ramachandran angles ($< 15^o$) are required to interconvert between the two bond types. There

is no disallowed energy region between the 3_{10}- and the α-conformation, and we therefore found a gradual transformation from one bond to the other [42, 65].

- The step to node 2 continues the 7-ring, but builds a new 3_{10}-bond. The energy difference of 3.63 kcal/mol should be mainly the effort of up-bending of the helix which has to exceed the gain of the bond forming.
- The step to node 3 dissolves the 7-ring and a 3_{10}-bond in residue 7. The large energy effort of 4.74 kcal/mol is consistent with this lost.
- For node 4 the MOLDEN reports the creation of the α-bond in residue 4. An energy gain of -0.61 kcal/mol seems to be in line with the formation of a new bond. Up to node 4 we should assume an ascent of the underlying RP.
- At node 6 the 3_{10}-bond of residue 8 is dissolved, and the corresponding 4_{13}-bond is enforced.
- At node 7 a 4_{13}-bond is new, this is consistent with the energy gain of -7.31 kcal/mol and further bending back of the helix.
- In node 8 two 3_{10}-bonds are exchanged, and a further 3_{10}-bond is exchanged with a 4_{13}-bond. Additionally, the funnel of the α-helix gives an energy gain for a further bending back of the helix, being consistent with -4.58 kcal/mol.
- Node 10 has to complete three 4_{13}-bonds, but they are already prepared in the preceding steps.
- The last step of the RP has to complete the last open 4_{13}-bond, and the straigtening of the helix.

Seen over all nodes, we state that the transformations along the RP do not contradict the estimated energy values of an H-bond, which are cited at the beginning of this paragraph. We can confirm that the transition back to the long α-helix part of the upright helix occurs through some 3_{10}-bond transients [33]. The treatment of the behaviour of the different H-bonds along the found reaction valley strengthens our statement that using only a constrained number of leading coordinates is too simple a model to make a reliable estimate of the true reaction pathway. Of course, H-bonds outside the hinge concern other degrees of freedom than the Ramachandran angles ϕ_8 and ψ_8.

The three non-H-bonded NH groups at the N-terminus of the upright α-helix give a net positive charge, and the three free CO groups give a negative charge at the C-terminus. The peptide planes are roughly parallel to the helix axis and the dipoles within the helix are aligned, i.e. all C=O groups point in the same direction and all N-H groups point the other way. Additionally, the H-bonds are almost parallel to the helix axis. Thus, an α-helix has a large dipole moment because the dipoles associated with the H-bonds align to form a macro-dipole [66]. The dipoles of the 3_{10}-helix are not so well aligned as in the α-helix, they are quite tilted. Consequently, the bend of this pentadecamer of alanine should have a much lower average dipole moment, and a measurement like in ref.[66] could decide between the structures.

Discussion of a turn of H-Cα against Cα-CH_3 in residue 8 of Ac(Ala)$_{15}$NHMe

The linear α-helix, **2**, is composed by L-enantiomers of alanine because the Me groups in the outer direction have a better stereochemical position than in the reverse case. However, the bent minimum structure of the helix **3**, see Fig. 6, is not the deepest minimum which is possible for such a bent structure. If in the 7-ring an exchange of the bonds H-Cα against Cα-Me is done, there is another minimum of the bent helix, see Fig. 10. The quasi-D-form has an energy of -3957.341 94 au, thus it is 1.224 kcal/mol lower than the L-form. Contrary, for the α-helix, we may determine the quasi-D-form of residue 8 only. It is a higher minimum at -3957.357 87 au, thus it is 1.505 kcal/mol only over the usual L-form. The reason may be that the methyl side chains in polyalanine make only minimal interactions with other side chains.

The former Fig. 7 shows the energy profile over a possible reaction path along the valley from the higher L-form of the bend to the lower L-form of the α-helix. The question arises, whether we can determine a pathway from the deepest minimum of the reactant side to the deepest minimum of the product side? However, the back-bending of the helix would now meet the additional conversation of the enantiomers. We do not know the minimal energy for this conversation, however, in the example of alanine dipeptide we had found a TS from L-form to quasi-D-form of the 7-ring of the C7ax conformation at $\approx$75 kcal/mol over the L-enantiomer. Approximating the transition structure of the L-form of the bent helix to the L-form of the α-helix by the energy difference to the highest node, node 6 in Fig. 7, of 13 kcal/mol, and guessing that the enantiomerisation of one residue of the helix needs approximately the analogous amount of the alanine dipeptide, we have to assume a very high ridge between two possible, but parallel reaction valleys, one valley for L to L-form of the back-bending, and another valley for bending back the quasi-D to quasi-D-form.

In Fig. 11 is shown the energy profile of a possible pathway from the lower quasi-D-form of the bent to the lower L-form of the α-helix. At the start, ϵ was set to 0.001 8 Å. The calculation was more complicated than the former calculation of Fig. 7. Often the corrector walks in contrary directions from node to node and in different dimensions. Though the ϵ threshold is tighter than in the former calculation, corresponding energy differences of some nodes are slightly higher. At node 4, ϵ was set to 0.002 5 Å, the value of the former calculation. Additionally, the corrector did not converge at node 7, but was cut externally after 83 iterations. The tracking for minima from higher nodes converges either into the quasi-D-form of the bend, or into the quasi-D-form of residue 8 of the α-helix. The step from node 8 to 9 seems to be too short: the corrector for node 9 walks "back" in the sink of the D-form. So, up to node 9, we are in the valley of a transition from quasi-D to quasi-D-form. Accidentally, the predictor step from node 9 to node 10 is large enough to do a step over the assumed high ridge of the enantiomerisation barrier. The predictor step, eq.(3), is a pure matter of the geometrical structure of the molecule. It does not depend from the PES of the molecule. The corrector of node 10 then converges to a point in the pocket of the L-form of the α-helix. From node 9 to node 10 we jump over the corresponding second TS. Summarising, the GS method is able to also determine a pathway of a complicate reorganisation of a large molecule.

However, one should not try to use the method for distances of reactant and product being too far. We should have in mind that the tool of a NT is not in every case a good model for a valley ground way [13, 18]. An exact NT can wander along the slope of the PES, though it will in every case pass the next stationary point. Figure 12 is an exaggerated model for two parallel valleys divided by a ridge in between. The model function is

$$\begin{aligned} pes(x,y) = 0.3\,Exp(-(x-0.1)^2-(y-0.1)^2) - 2\,Exp(-0.9\,(x-1.3)^2 - 0.3\,(y+1.6)^2) \\ -2\,Exp(-(x+1.5)^2-(y+1.7)^2) - 2\,Exp(-0.9\,(x-1.4)^2-(y-1.8)^2) \\ -2\,Exp(-(x+1.3)^2 - 0.25\,(y-1.23)^2)\ . \end{aligned} \quad (5)$$

The minima top right and bottom left are the two lowest minima. An NT is shown (the fat dashed curve) which uses the direction, **r**, between the lowest minima for the search direction. The NT starts at the lowest minima in this direction, **r**, but then it has to go off from the "good" direction to find the next saddle point which is the SP of the corresponding valley. The NT does not go along the shortest pathway in this case, over the summit, a saddle of index 2 in the centre of the Figure. Note that in the constructed case of Fig. 12, the NT to direction **r** does not connect the two lowest minima, at all. In the praxis of the GS method we do circumvent such a situation by an adapted, variable search direction between the last but one node and the final product structure [9]. Thus, we change the special NT from node to node which we employ for the RP. Of course, there is a family of NTs between the two minima at the bottom of Fig. 12 which connect the minima. One member of the family is the given thin dashed curve.

Using that exaggerated model of Fig. 12 for our RP problem in polyalanine, and assuming the left valley for the L-form, and the right valley for the D-form, we may better understand the result of our calculation. A pathway in direction of the y-axis is an RP inside of one enantiomeric form: it is relatively easy to calculate the NT using the y-axis for the search direction. However, if we chose the diagonal y=x direction between the two deepest minima, the corresponding NT is more complicated, and its numeric determination should also be more difficult than the pure valley ground path. That is what we have observed for the profile of Fig. 11.

Conclusion

The growing string method with Newton trajectories introduced as a global search method [9] is proposed to be applied to finding a TS region between a reactant and a product. The present method can find the TS region without an initial guess. This study demonstrates the potential of the GS-NT method in combination with the Berny method for the SP search in conformational transitions in relatively high-dimensional problems: The performance of the GS-NT method was demonstrated for an isomerisation and an enantiomerisation of alanine dipeptide, as well as for a bending of a medium polyalanine helix. For the enantiomerisation, we employed a big number of nodes to catch the jump of a single H atom in a large molecular remainder. For the hinge-bending motion of the polyalanine, we had to use a coarse-grained approach of the RP by only 10 nodes due to the computational costs. The predictor of the method acts globally, where the corrector of the method finds the reaction coordinate by a quasi local search, see Figs. 1 and 8. The relatively high dimension of the example of an unsolved chain of the model peptide Ac-(Ala)$_{15}$-NHMe, with dimension 486, demonstrates that the method still works in such a high-dimensional case, though the projection of the gradient in eq.(1) by the search direction, **r**, concerns "only" one dimension. In spite of the observations that folding is an inherently complex process, we are able to use a one-dimensional RP. The approximation of the reaction path is given numerically with the help of a Newton trajectory, which is another reaction path definition than the usually used IRC, compare also the result of the nudged elastic band method in Fig. 2 of ref. [67] for alanine dipeptide. It is important to note that in the case of polyalanine no TS has been located. But incidentally, we have to report for the first time the formation of a stable bend of a 15-mer helix with a 2_7-bond at the hinge.

On the other hand, we guess that a non-global search method like the local following of an NT [13] by its tangent, or the direct following of a valley pathway by a gradient extremal [15], may go into troubles on a peptide PES. The reason is the overwelming number of configurations in the sink pockets of some kernel structures [49, 68]. Between near configurations a large number of flat valley pathways of no chemical interest may exist which have to be followed by purely local methods. In contrast, the global leading line of the predictor of the GS method uses current node and product of the reaction of interest. Large predictor steps can jump over many flat wells. The locally working corrector usually finds approximations of the nodes of the global reaction path between the two minima.

For low dimensions is the GS-NT method comparable to other methods in efficiency, cf.[2] for the example alanine dipeptide. However, in paper [2], the authors did not prove the downhill walking on the other side from the TS to the product region. The calculation of only "half" of the RP of the isomerisation explains (in part) the low number of reported gradient calls [2]. We find that the complete calculation of the RP is quite useful for the illustration of the reaction in question, especially for unknown reactions.

A natural drawback of the GS method is the successive calculation of the RP nodes: the calculation of all nodes together is not parallelisable. (It would be possible for NTs also by projection of an overall predictor line, see ref. [8]. However, that is not a growing string method.) In the GS method, every new step starts when the previous step has been finished. However, if the foregoing step finishes near the global RP, the next predictor will be not so far away from the RP, and the next corrector will only have to do a moderate computational task.

A method producing the reaction path of the potential energy surface (PES) for a given chemical composition and its possible reactions was proposed and applied. The present growing string method with Newton trajectories (GS-NT) can be used in combination with the Berny method to discover a transition state (TS), which is connected with two equilibrium structures (EQ). Different further TSs belonging to the given reactant and product may be detectable by different search directions, **r**. Starting from a TS and using conventional downhill-walk techniques, another EQ may also be found. After having found other EQs, or using other known EQs, the present method can be used again to find out networks of reaction pathways around the new EQs. In this manner, the topographic map of the PES can be explored, cf. [13, 19, 49]. Such procedures may be used to disclose the main reaction pathways to chemical products starting from a given chemical composition.

Acknowledgement: We could use computer time of Leipziger Universitäts-Rechenzentrum, on clusters of the type Hewlett Packard Itanium2. I am also grateful to the reviewers of this paper for their careful reviews and wonderful suggestions.

References

[1] Peters, A.; Heyden, A.; Bell, A.T.; Chakraborty, A. J Chem Phys, 2004, 120, 7877.

[2] Maeda, S.; Ohno, K. Chem Phys Lett, 2005, 404, 95.

[3] Schlegel, H. B. J Comput Chem, 2003, 24, 1514.

[4] Henkelman, G.; Jonsson, K. J Chem Phys, 2000, 113, 9978.

[5] Trygubenko, S. A.; Wales, D. J. J Chem Phys, 2004, 120, 2082.

[6] E, W.; Ren, W.; Vanden-Eijnden, E. Phys Rev B, 2002, 66, 052301;

[7] Crehuet, R.; Field, M.J. J Chem Phys, 2003, 118, 9563.

[8] Quapp, W. J Computat Chem, 2004, 25, 1277.

[9] Quapp, W. J Chem Phys, 2005, 122, 174106.

[10] Cerjan, C. J.; Miller, W. H. J Chem Phys, 1981, 75, 2800.

[11] Frisch, M. J.; et al. Gaussian03, Revision C.02, Gaussian, Inc, Wallingford CT, 2004.

[12] Hoffman, D. K.; Nord, R. S.; Ruedenberg, K. Theoret Chim Acta, 1986, 69, 265.

[13] Quapp, W.; Hirsch, M.; Imig, O.; Heidrich, D. J Computat Chem, 1998, 19, 1087;
Quapp, W. J Computat Chem, 2001, 22, 537.

[14] Anglada, J. M.; Besulu, E.; Bofill, J. M.; Crehuet, R. J Computat Chem, 2001, 22, 387;
Bofill, J. M.; Anglada, J. M. Theor Chem Acc, 2001, 105, 463;
Crehuet, R.; Bofill, J. M.; Anglada, J. M. Theor Chem Acc, 2002, 107, 130.

[15] Quapp, W.; Hirsch, M.; Heidrich, D. Theoret Chem Acc, 2000, 105, 145.

[16] Wang, S.; Smith, S. C. Chem Phys, 2006, 326, 204.

[17] Hirsch, M.; Quapp, W. J Molec Struct, THEOCHEM, 2004, 683, 1.

[18] Quapp, W. J Computat Chem, 2001, 22, 537.

[19] Carr, J. M.; Trygubenko, S. A.; Wales, D. A. J Chem Phys, 2005, 122, 234903;
Carr, J. M.; Wales, D. A. J Chem Phys, 2005, 123, 234901;

[20] Beachy, M. D.; Chesman, D.; Murphy, R. B.; Halgreen, T. A.; Friesinger, R. A.
J Am Chem Soc, 1997, 119, 5908.

[21] Liu, D.; Nocedal, J. Math Programm, 1989, 45, 503.

[22] Berger, R.; Quack, M. Chem Phys Chem, 2000, 1, 57;
Laerdahl, J. K.; Wesendrup, R.; Schwerdtfeger, P. Chem Phys Chem., 2000, 1, 60.

[23] Chass, G. A.; Sahai, M. A.; Law, J. M. S.; Lovas, S.;Farkas, Ö.; Perczel, A.; Rivail, J.-L.;
Csizmadia, I.G. Int J Quant Chem, 2002, 90, 933.

[24] Bai, F.; Wang, W.; Min, W. preprint, 2002, http://arxiv.org/abs/physics/0209069;
Wang, W.; Min, W.; Gong, Y. Acta Phys-Chim Sin, 2005, 21, 1186.

[25] Levinthal, C. in: Mössbauer Spectroscopy in Biological Systems, Proc. Meeting at Allerton House, Monticello, Illinois, (eds DeBrunner, P.; Tsiberis, J.; Munck, E.) Uni Illinois Press, Urbana, 1962, p.22.

[26] Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Proteins, 1995, 21, 167.

[27] Bai, Y. Biochem Biophys Res Comm, 2003, 305, 785.

[28] Kennedy, R. J.; Walker, S. M.; Kemp, D. S. J Am Chem Soc, 2005, 127, 16961.

[29] Ramachandran, G. N.; Sasisekharan, V. Adv Protein Chem, 1968, 23, 283.

[30] Doig, A. J. Biophys Chem, 2002, 101-102, 281.

[31] Schaftenaar, G. www.cmbi.ru.nl/molden

[32] Ösapy, K. ; Young, W. S.; Bashford, D.; Brooks III, C. L.; Case, D. A. J Phys Chem, 1996, 100, 2698.

[33] Young, W. S.; Brooks III, C. L. J Mol Biol, 1996, 259, 560.

[34] Fiori, W. R.; Miick, S. M.; Millhauser, G. L. Biochem, 1993, 32, 11957.

[35] Millhauser, G. L.; Stenland, C. J.; Hanson, P.; Bolin, K. A. van de Ven, F. J. M. J Mol Biol, 1997, 267, 963.

[36] Alonso, J. L.; Echenique, P. J Comput Chem, 2005, 27, 238.

[37] Wieczorek, R.; Dannenberg, J. J. J Am Chem Soc, 2004, 126, 14198.

[38] Okamoto, Y.; Hansmann, U. E. J Phys Chem, 1995, 99, 11276.

[39] Sali, A.; Shakhnovich, E. I.; Karplus, M. Nature, 1994, 369, 248.

[40] Bryngelson, J.; Wolynes, P. J Chem Phys, 1989, 93, 6902.

[41] Quapp, W. www.math.uni-leipzig.de/∼quapp/ALA15

[42] Rohl, C. A.; Doig, A. J. Protein Sci, 1996, 5, 1687.

[43] Wang, C.-S.; Zhang, Y.; Gao, K.; Yang, Z.-Z. J Chem Phys, 2005, 123, 24307.

[44] Bertsch, R. A.; Vaidehi, N.; Chan, S. I.; Goddard III, W. A. Proteins, 1998, 33, 343.

[45] Podtelezhnikov, A. A.; Wild, D. L. Proteins, 2005, 61, 94.

[46] Wang, Y.; Kuczera, K. J Phys Chem, 1997, 101, 5205.

[47] Morozov, A. V.; Tsemekhman, K.; Baker, D. J Phys Chem B, 2006, 110, 4503.

[48] Koskowski, F.; Hartke, B. J Comput Chem, 2005, 26, 1169.

[49] Mortensen, P. N.; Evans, D. A.; Wales, D. J. J Chem Phys, 2002, 117, 1363.

[50] Takano, M.; Takahashi, T.; Nagayama, K. Phys Rev Lett, 1998, 80, 5691.

[51] Evans, D. A.; Wales, D. J. J Chem Phys, 2003, 118, 3891.

[52] Mortensen, P. N.; Wales, D. J. J Chem Phys, 2001, 114, 6443.

[53] Komatsuzaki T.; Hoshino K.; Matsunaga Y.; Rylance G. J.; Johnston R. L.; Wales, D. J. J Chem Phys, 2005, 122, 084714.

[54] Rhee, Y. M.; Pande, V. S. Chem Phys, 2006, 323, 66;
Petrone, P.; Pande, V. S. Biophys J, 2006, 90, 1583.

[55] Gruebele, M. Curr Opinion Struct Biol, 2002, 12, 161.

[56] Snow, C. D.; Rhee, Y. M.; Pande, V. S. Biophys J, 2006, 91, 14.

[57] Ren, W.; Vanden-Eijnden, E.; Maragakis P.; E, W. J Chem Phys, 2005, 123, 134109.

[58] Brooks, C. L. J Phys Chem, 1996, 100, 2546.

[59] Pauling, L.; Corey, R. B.; Branson, H. R. Proc Natl Acad Sci USA, 1951, 37, 205.

[60] Shen, S.-Y.; Yang, D.-Y.; Selzle, H. L.; Schlag, E. W.
Proc Natl Acad Sci USA, 2003, 100, 12683.

[61] Lee, I.-H.; Kim, S.-Y.; Lee, J. Chem Phys Lett, 2005, 412, 307.

[62] Baker, J. J Chem Phys, 2006, 125, 014103.

[63] Vijayakumar, S.; Vishveshwara, S.; Ravishanker, G.; Beveridge, D. L. in "Modeling the Hydrogen Bond", ACS Symposium Series, 1994, **569**, 175.

[64] Ohkubo, Y. Z.; Brooks, C. L. Proc Natl Acad Sci USA, 2003, 100, 13916.

[65] Toniolo, C.; Benedetti, E. Trends Biochem Sci, 1991, 16, 350.

[66] Dugourd, P.; Antoine, R.; Breaux, G.; Broyer, M.; Jarrold, M. F.
J Am Chem Soc, 2005, 127, 4657.

[67] Chu, J.-W.; Trout, B. L.; Brooks, B. R. J Chem Phys, 2003, 119, 12708.

[68] Vila, J. A.; Ripoll, D. R.; Scheraga, H. A. Proc Natl Acad Sci USA, 2000, 97, 13075.

Table I Torsional angles of two minimum structures of Ac-(Ala)$_{15}$NHMe: linear and bent helix. (The torsionals are only given if the difference is larger than 10 degrees).

dih[1)]	type[2)]	linear	bent helix	|dif|
20 (4)	χ	174.896	162.496	12.4
22 (4)	ϕ	-60.957	-73.178	12.221
28 (5)	ψ	-44.925	-33.987	10.938
31 (5)	Θ	136.112	146.357	10.245
33 (6)	ψ	-44.07	-18.594	25.476
35 (7)	χ	175.146	142.353	32.793
36 (6)	Θ	137.203	161.964	24.761
37 (7)	ϕ	-60.988	-92.798	31.81
38 (7)	ψ	-44.549	5.167	49.716
40 (8)	χ	-185.425	-57.389	128.036
41 (7)	Θ	136.663	185.144	48.481
42 (8)	ϕ	-61.68	70.756	132.436
43 (8)	ψ	-42.981	-58.78	15.799
46 (8)	Θ	138.417	121.555	16.862
50 (10)	χ	173.842	160.386	13.456
52 (10)	ϕ	-62.345	-73.99	11.645
53 (10)	ψ	-43.603	-15.063	28.54
55 (11)	χ	176.228	151.023	25.205
56 (10)	Θ	137.371	164.851	27.48
57 (11)	ϕ	-59.886	-84.813	24.927
100 (4)	H	54.053	41.457	12.596
115 (7)	H	54.183	22.38	31.803
120 (8)	H	53.677	184.045	130.368
130 (10)	H	53.015	40.822	12.193
135 (11)	H	55.118	29.469	25.649
160 Me	Me	126.617	166.052	39.435
161 Me	Me	6.636	45.808	39.172
162 Me	Me	-112.978	-73.673	39.305
energy[3)]	a.u.	-3957.360 27	-3957.339 97	0.020 30

[1)]Dihedral angles in a z-matrix representation, in the bracket (..) is the residue. Full geometry parameters are available on web-page [41] / supporting information.
[2)]Backbone Torsionals (Ramachandran coordinates) are: ϕ is C′-N-Cα-C′, ψ is N-Cα-C′-N, ω is Cα-C′-N-Cα, others: χ is C-Cα-N-C′ (of the carbon of the side -CH_3), Θ is O-C′-Cα-N, H is Hα-Cα-N-C′, and Me is H-C-N-C′ of the CH_3 of the cap of the C-terminus.
[3)]Energy of a B3LYP/6-31G calculation.

Table II H-bonds of Ac-(Ala)$_{15}$-NHMe along the reaction coordinate.

bond	bent helix 1)	node2 2)	node3	node4	node6	node7	node8	node9	node10	lin.helix
3_{10} 3)	2-6⋯94-13	∃ 4)	∃	∃	∃	∃	∃	∃	∃	2-6⋯94-13
4_{13} 3)	-	-	-	*N*	∃	∃	∃	∃	∃	2-6⋯99-18
4_{13}	7-11⋯104-23	∃	∃	∃	∃	∃	∃	∃	∃	7-11⋯104-23
4_{13}	12-16⋯109-28	∃	∃	∃	∃	∃	∃	∃	∃	12-16⋯109-28
4_{13}	-	-	-	-	-	*N*	∃	∃	∃	17-21⋯114-33
3_{10}	22-26⋯114-33	∃	*S*	-	-	-	-	-	-	-
4_{13}	-	-	-	-	-	-	-	-	*N*	22-26⋯119-38
3_{10}	27-31⋯119-38	∃	∃	∃	∃	∃	*S*	-	-	-
4_{13}	-	-	-	-	-	-	-	-	*N*	27-31⋯124-43
3_{10}	-	-	-	-	-	-	36⋯124	*S*	-	-
4_{13}	-	-	-	-	-	-	-	-	*N*	32-36⋯129-48
3_{10}	-	41⋯129	∃	∃	∃	∃	*S*	-	-	-
4_{13}	-	-	-	-	-	-	*N*	∃	∃	37-41⋯134-53
2_7 5)	39-38-37-41	*S*	-	-	-	-	-	-	-	-
	⋯124-43-42	-	-	-	-	-	-	-	-	-
3_{10} 3)	42-46⋯134-53	*S*	-	-	-	-	-	-	-	-
4_{13} 3)	42-46⋯139-58	∃	∃	∃	∃	∃	∃	∃	∃	42-46⋯139-58
4_{13}	47-51⋯144-63	∃	∃	∃	∃	∃	∃	∃	∃	47-51⋯144-63
4_{13}	-	-	-	-	-	-	-	-	-	52-56⋯149-68
3_{10}	57-61⋯149-68	∃	∃	∃	*S*	-	-	-	-	-
3_{10}	62-66⋯154-73	∃	∃	∃	∃	∃	∃	∃	∃	62-66⋯154-73
3_{10}	67-71⋯159-78	∃	∃	∃	∃	∃	∃	∃	∃	67-71⋯159-78

1) Numbers of atoms in a z-matrix representation in ascending order:
the sequence is C′O⋯HN

2) Nodes 1 and 5 are omitted;
(node 1 corresponds to bent helix, node 5 corresponds to node 4)

3) Hybrid α / 3_{10} - helical structure

4) ∃: the bond continues to exist, *S*: the bond is coming off, *N*: the bond is new
(reported using MOLDEN4.4 [31])

5) The 7-membered ring in the bend is given in the sequence CαNC′O⋯HNC′

Figure captions

Fig. 1 Approximation of a Newton trajectory by the GS method for alanine dipeptide between the minimum $C_{7\,ax}$ (start) and minimum C5 (finish). Method: modified CG+ optimisation [9] at ϵ=0.003 5. Two coordinates (ϕ , ψ) are drawn in a Ramachandran diagram where (in the Figure) the PES is optimised for all other coordinates at a lattice of points. 9 nodes are used for the string. Predictor steps are large, corrector steps are usually very small. The curvilinear line is a fit through the calculated chain of nodes. The SP is determined by Berny optimisation [11] starting at the highest points of the string.

Fig. 2 Approximation of the energy profile (bullets) of a Newton trajectory by the GS method for alanine dipeptide between the minima $C_{7\,ax}$ and C5, see Fig.1. 9 nodes are used for the string. Included are energies of predictor steps and of adjacent corrector loops. Energy in kcal/mol.

Fig. 3 Numbering of the alanine dipeptide proposed by Chass et al.[23].

Fig. 4 Approximation of the energy profile (points) of a Newton trajectory by the GS method for alanine dipeptide between the minima of the L-enantiomer of $C_{7\,ax}$ (left) and the corresponding D-form (right). Bullets are minima, stars are saddle points. 30 nodes are used for the string from left L-enantiomer to next Min. Then 8 nodes are used for the string to the next, enantiomeric Min, and the pathway from the D-enantiomer to that Min is plotted in reverse order. It has also 30 nodes. The structure top left is the transition state of the moving H atom in the L-form. The structure top right is the D-form minimum where the moving H atom has reached the target of the N atom. The structure at bottom centre is the flat transition intermediate with σ_v symmetry between two real enantiomeric minima.

Fig. 5 Approximation of the energy profile (points) of Newton trajectories by the GS method for alanine dipeptide between the minima of the L-enantiomer of $C_{7\,ax}$ (left) and the corresponding D-form (right) using O10 for a parking place of H11. Bullets are minima, stars are saddle points. 30 nodes are used for the strings. The pathway from the D-enantiomer to central minimum is plotted in reverse order. The structure top right is the TS H11 at atom C9, the structure top left is the TS H11 at O10 where the corresponding minimum is bottom centre.

Fig. 6 All-atom representation of Ac-(Ala)$_{15}$-NHMe: mainly α-helix (left) and bent minimum structure (right). Dashes are H-bridges. Side chains are Me groups sitting at the Cα atoms. They point outward from helix axis and are generally oriented towards its N-terminus. The α-helix has $\approx$3.6 residues per turn [34], so the side -CH_3 groups do not directly overlap from i to $i+4$, and the α-bonds are a little skew. The 3_{10}-bonds, in contrast, are the skew bonds in this structure. The hinge of the bent helix is a 7-ring. Atoms are rendered using colors: carbon, orange; nitrogen, blue; oxygen, red; hydrogen, white. Structures are processed with Molden4.4 software [31].

Fig. 7 Energy profile of an approximation of a Newton trajectory by the GS method for an isomerisation from the minimum of the bent helix (B) of Ac-(Ala)$_{15}$-NHMe to the minimum of the linear helix (L), see Fig. 6. Symbol ?? could be the location of the TS, see subsection search for a TS. Energy in kcal/mol.

Fig. 8 Ramachandran angles ϕ_8 and ψ_8 for the corrector loops (bullets) under the approximation of the Newton trajectory of Ac-(Ala)$_{15}$-NHMe. Numbers are the nodes of Fig. 7. The thin line is the projection of the NT in the plane (ϕ_8,ψ_8).

Fig. 9 Eight unsuccessful examples of a search for a TS on the RP of Ac-(Ala)$_{15}$-NHMe. Start point is a linear combination on the pathway from a bent structure to α-helix between nodes 4 and 5, see text. Initial energy is at the top. The energy of every step is set in relation to the leading coordinate, $\phi 8$. The energy scale is in kcal/mol over the α-helix. The first $\approx$100 steps of every search curve are shown. The calculations fall into the funnels of the bend or of the α-helix under meandering. But some curves pass over others. The inlay conformation with $\phi 8$=16^o is displayed as wire frame.

Fig. 10 View onto the hinge, the 7-ring of Ac-(Ala)$_{15}$-NHMe in bent structure. On the left hand side is the L-enantiomer, c.f. Fig. 6, on the right is the quasi-D-form.

Fig. 11 Energy profile of two approximations of Newton trajectories by the GS method for an isomerisation of the bent helix of Ac-(Ala)$_{15}$-NHMe to the linear helix. The string with point nodes is the pathway from quasi-D-form of the bend to L-form of the α-helix, see text. The other string is for comparison that one from the former Fig. 7. The step T is the transition from D-valley into L-valley.

Fig. 12 2D model PES for two parallel valleys as a counter example: the dashed curve is the Newton trajectory for a transition which uses the direction between the deepest minima, M_1 and M_2, for the search direction. The NT is not well adapted to the valley floor lines, see text. Thin curves are the NTs in axes directions: they are good valley lines.

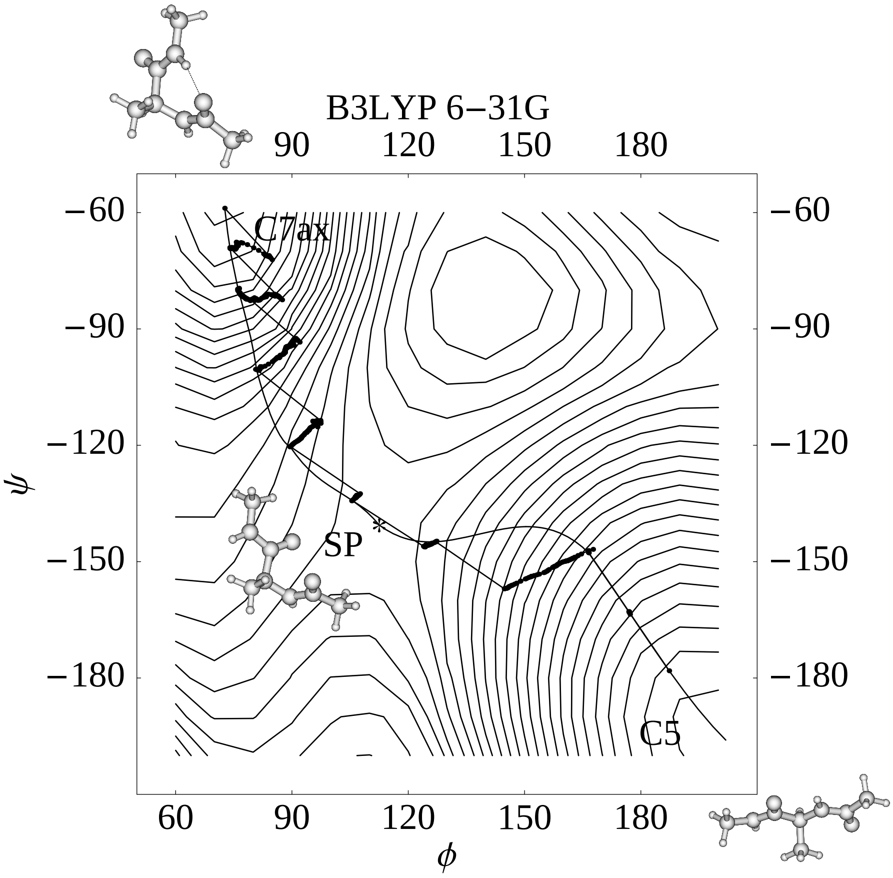


Figure 1: Approximation of a Newton trajectory by the GS method for alanine dipeptide between the minimum $C_{7\,\mathrm{ax}}$ (start) and minimum C5 (finish). Method: modified CG+ optimisation [9] at ϵ=0.003 5. Two coordinates (ϕ , ψ) are drawn in a Ramachandran diagram where (in the Figure) the PES is optimised for all other coordinates at a lattice of points. 9 nodes are used for the string. Predictor steps are large, corrector steps are usually very small. The curvilinear line is a fit through the calculated chain of nodes. The SP is determined by Berny optimisation [11] starting at the highest points of the string.

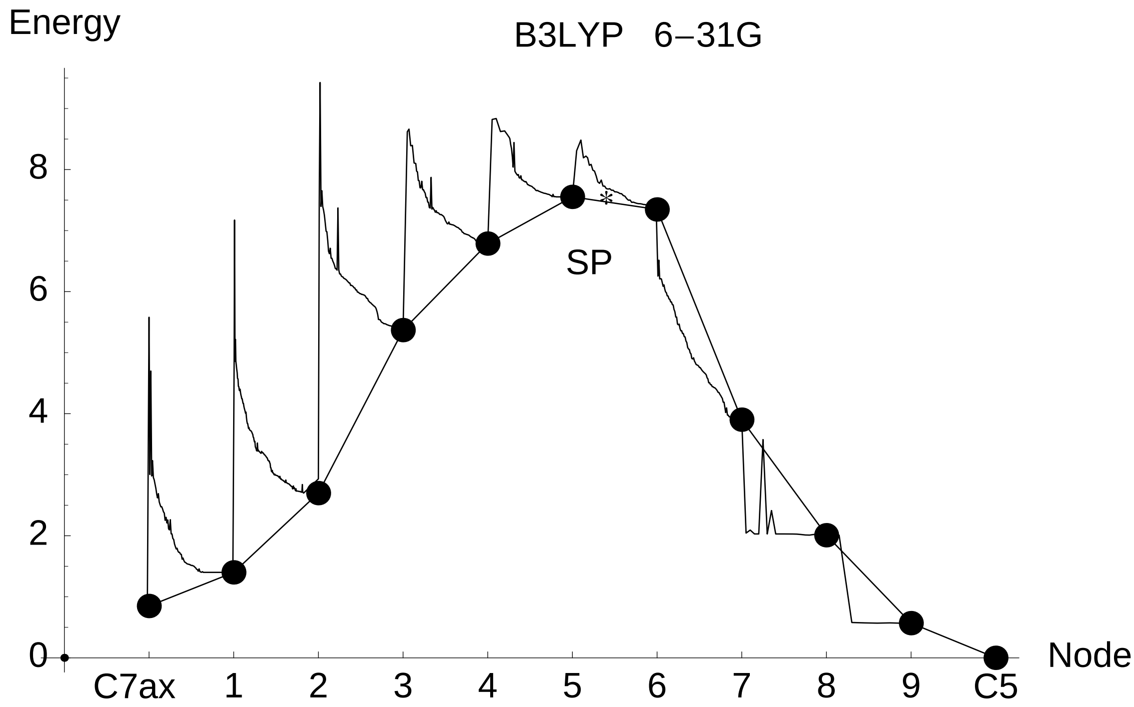


Figure 2: Approximation of the energy profile (bullets) of a Newton trajectory by the GS method for alanine dipeptide between the minima $C_{7\,ax}$ and C5, see Fig.1. 9 nodes are used for the string. Included are energies of predictor steps and of adjacent corrector loops. Energy in kcal/mol.

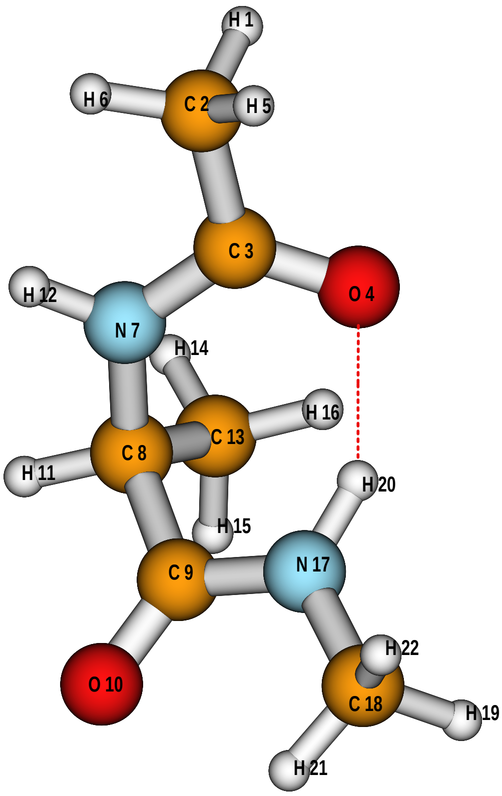


Figure 3: Numbering of the alanine dipeptide proposed by Chass et al.[23].

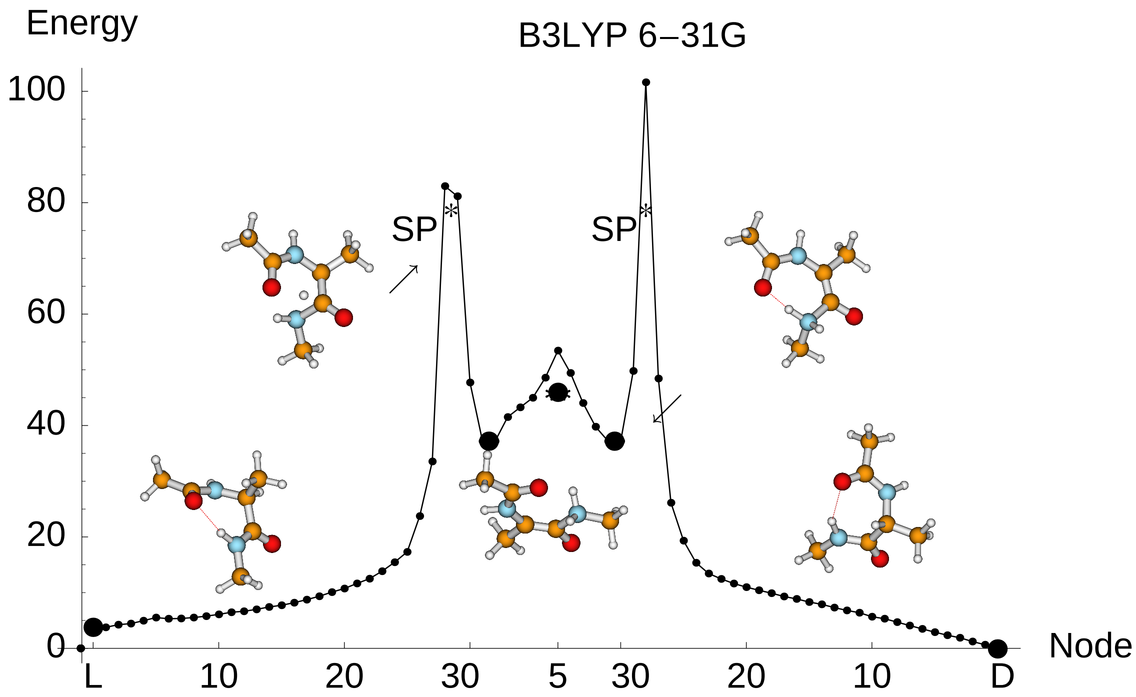


Figure 4: Approximation of the energy profile (points) of a Newton trajectory by the GS method for alanine dipeptide between the minima of the L-enantiomer of $C_{7\,ax}$ (left) and the corresponding D-form (right). Bullets are minima, stars are saddle points. 30 nodes are used for the string from left L-enantiomer to next Min. Then 8 nodes are used for the string to the next, enantiomeric Min, and th e pathway from the D-enantiomer to that Min is plotted in reverse order. It has also 30 nodes. The structure top left is the transition state of the moving H atom in the L-form. The structure top right is the D-form minimum where the moving H atom has reached the target of the N atom. The structure at bottom centre is the flat transition intermediate with σ_v symmetry between two real enantiomeric minima.

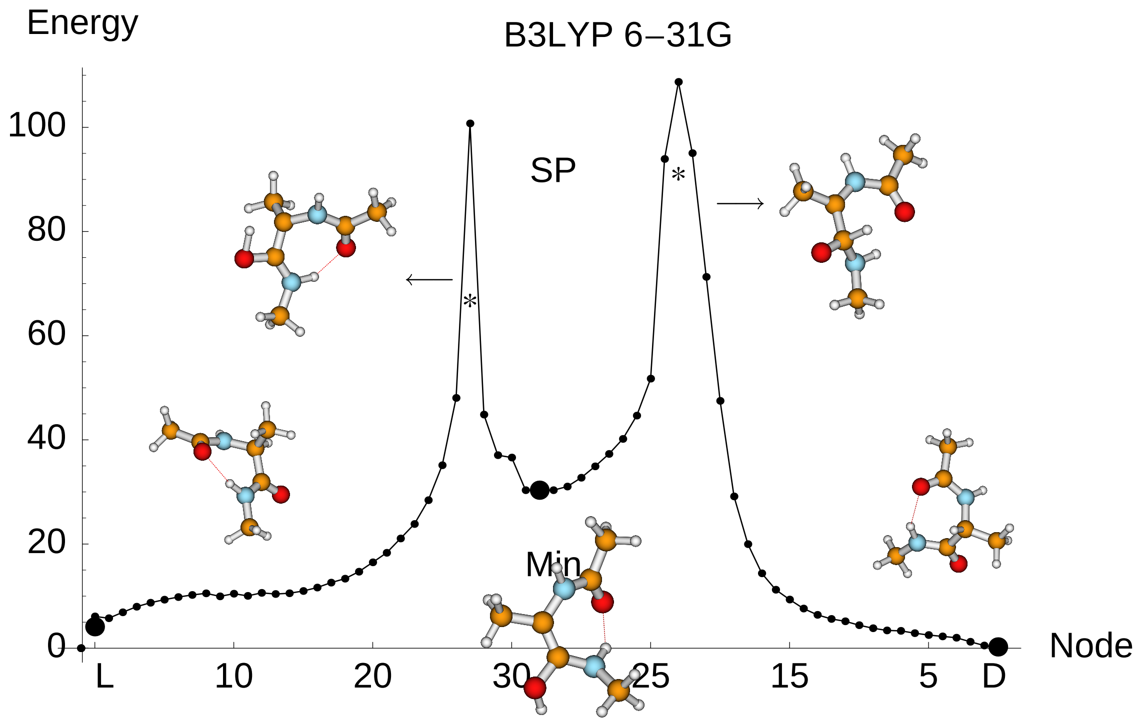
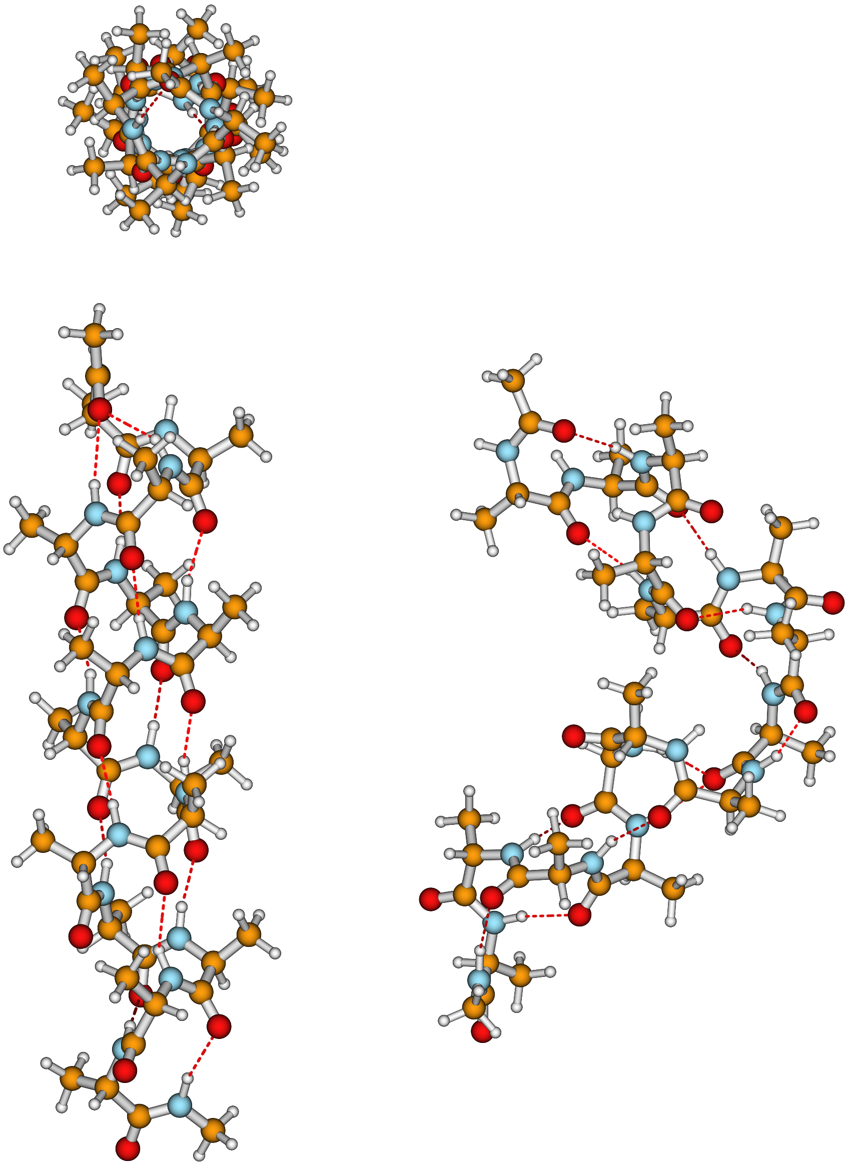


Figure 5: Approximation of the energy profile (points) of Newton trajectories by the GS method for alanine dipeptide between the minima of the L-enantiomer of $C_{7\,\mathrm{ax}}$ (left) and the corresponding D-form (right) using O10 for a parking place of H11. Bullets are minima, stars are saddle points. 30 nodes are used for the strings. The pathway from the D-enantiomer to central minimum is plotted in reverse order. The structure top right is the TS H11 at atom C9, the structure top left is the TS H11 at O10 where the corresponding minimum is bottom centre.



2

3

Figure 6: All-atom representation of Ac-(Ala)$_{15}$-NHMe: mainly α-helix (left) and bent minimum structure (right). Dashes are H-bridges. Side chains are Me groups sitting at the Cα atoms. They point outward from helix axis and are generally oriented towards its N-terminus. The α-helix has $\approx$3.6 residues per turn [34], so the side -CH_3 groups do not directly overlap from i to $i+4$, and the α-bonds are a little skew. The 3_{10}-bonds, in contrast, are the skew bonds in this structure. The hinge of the bent helix is a 7-ring. Atoms are rendered using colors: carbon, orange; nitrogen, blue; oxygen, red; hydrogen, white. Structures are processed with Molden4.4 software [31].

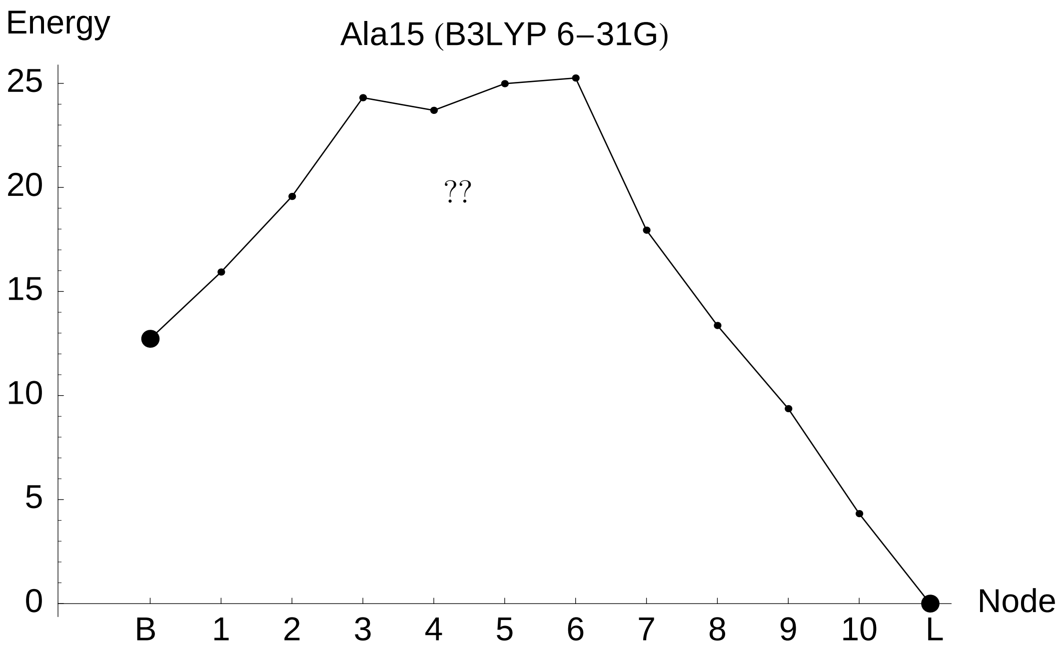


Figure 7: Energy profile of an approximation of a Newton trajectory by the GS method for an isomerisation from the minimum of the bent helix (B) of Ac-(Ala)$_{15}$-NHMe to the minimum of the linear helix (L), see Fig. 6. The question-marks ?? could be the location of the TS, see the guess in subsection search of a TS. Energy in kcal/mol.

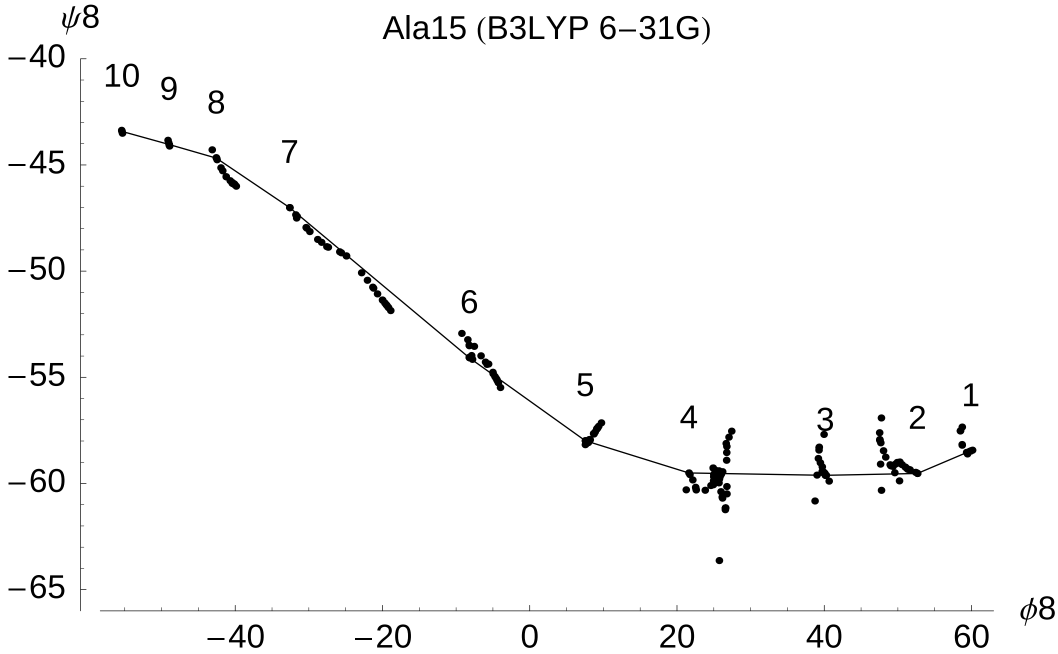


Figure 8: Ramachandran angles ϕ_8 and ψ_8 for the corrector loops (bullets) under the approximation of the Newton trajectory of Ac-(Ala)$_{15}$-NHMe. Numbers are the nodes of Fig. 7. The thin line is the projection of the NT in the plane (ϕ_8,ψ_8).

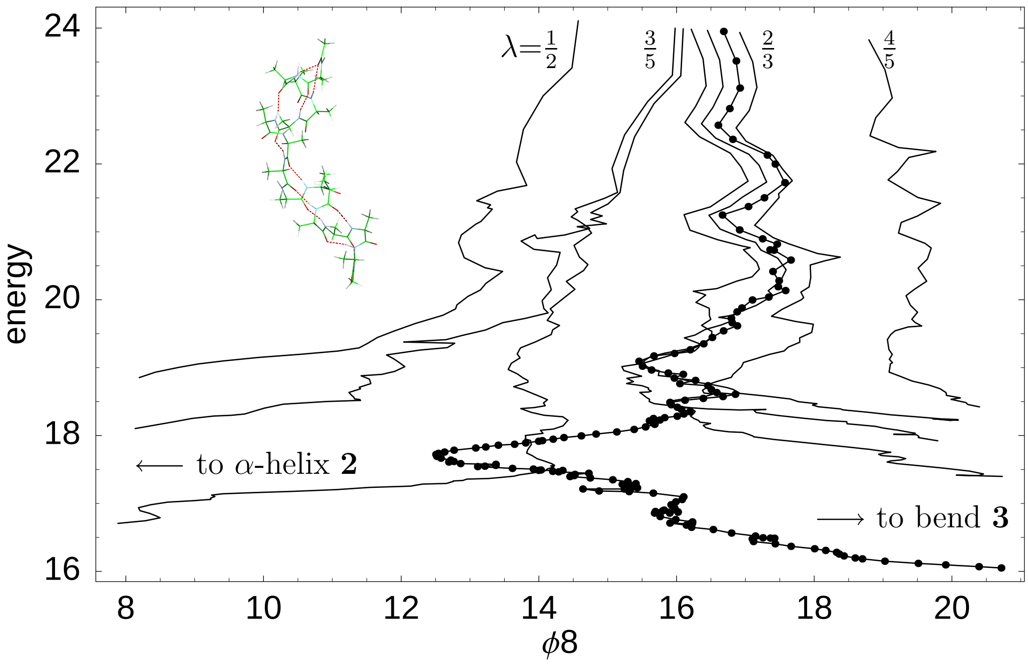
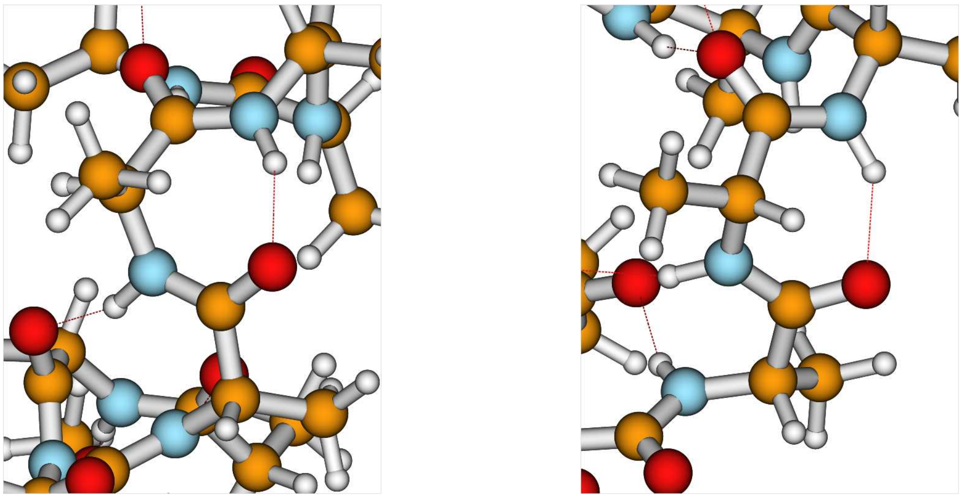


Figure 9: Eight unsuccessful examples of a search for a TS on the RP of Ac-$(Ala)_{15}$-NHMe. Start point is a linear combination on the pathway from a bent structure to α-helix between nodes 4 and 5, see text. Initial energy is at the top. The energy of every step is set in relation to the leading coordinate, $\phi8$. The energy scale is in kcal/mol over the α-helix. The first $\approx$100 steps of every search curve are shown. The calculations fall into the funnels of the bend or of the α-helix under meandering. But some curves pass over others. The inlay conformation with $\phi8$=16^o is displayed as wire frame.



L **D**

Figure 10: View onto the hinge, the 7-ring of Ac-(Ala)$_{15}$-NHMe in bent structure. On the left hand side is the L-enantiomer, c.f. Fig. 6, on the right is the quasi-D-form.

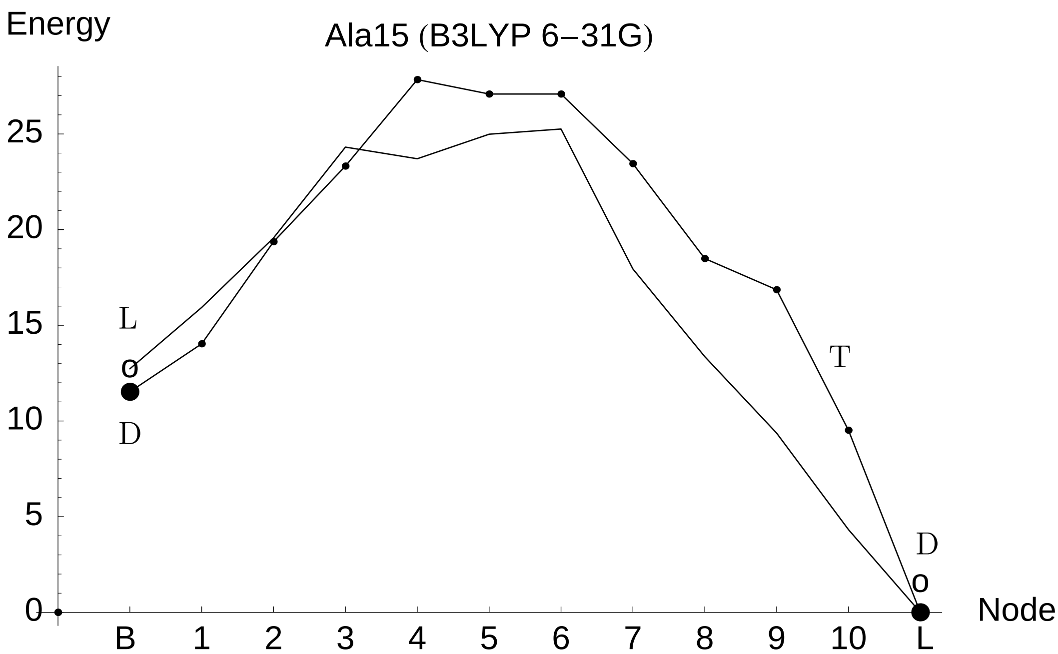


Figure 11: Energy profile of two approximations of Newton trajectories by the GS method for an isomerisation of the bent helix of Ac-(Ala)$_{15}$-NHMe to the linear helix. The string with point nodes is the pathway from quasi-D-form of the bend to L-form of the α-helix, see text. The other string is for comparison that one from the former Fig. 7. The step T is the transition from D-valley into L-valley.

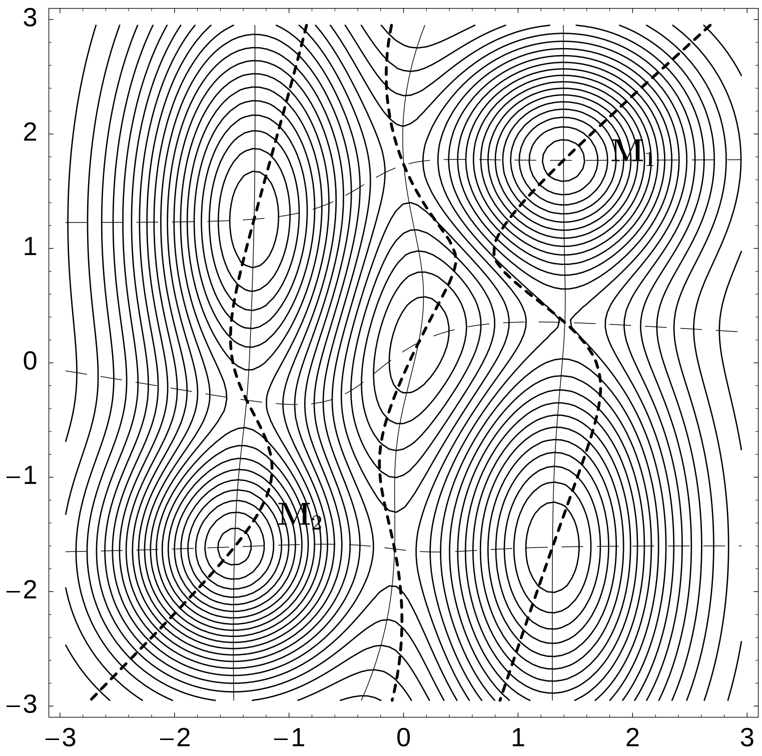


Figure 12: 2D model PES for two parallel valleys as a counter example: the dashed curve is the Newton trajectory for a transition which uses the direction between the deepest minima, M_1 and M_2, for the search direction. The NT is not well adapted to the valley floor lines, see text. Thin curves are the NTs in axes directions: they are good valley lines.