PROTEINS: Structure, Function, and Genetics 7:366–377 (1990)

The Building of Protein Structures From α-Carbon Coordinates

Paul E. Correa
The Procter & Gamble Company, Miami Valley Laboratories, P.O. Box 398707, Cincinnati, Ohio 45239-8707

ABSTRACT **A procedure for the construction of complete protein structures from only α-carbon coordinates is described. This involves building the backbone by sequential addition of Pro, Gly, or Ala residues. This main chain structure is then refined using molecular dynamics. Side chains are constructed by sequential addition of atoms with intermediate molecular dynamics refinement. For α lytic protease (a structure that is mostly β sheet) a backbone root mean square deviation (RMSD) of 0.19 Å and an overall RMSD of 1.24 Å from the crystallographic coordinates are attained. For troponin C (67% α-helix), where the coordinates are available only for the α-carbons, a backbone RMSD of 0.41 Å and an overall RMSD of 1.68 Å are attained (fits kindly provided by Dr. Michael James and Natalie Strynadka). For flavodoxin a backbone RMSD of 0.49 Å and an overall RMSD of 1.64 Å were attained.**

Key words: computer modeling, protein, structure, α-carbons

INTRODUCTION

The use of protein crystal structures for understanding biochemical processes has increased as more structures have become available, primarily through the compilation in the Brookhaven Protein Data Bank.[1] However, many of these data entries contain only the α-carbon coordinates. This is the result of many factors, including the need for further refinement of the coordinates before publication. This presents a frustrating situation for a worker who could use even a crude complete structure, but cannot use the incomplete structure. In addition, a map of the α-carbons is usually generated early in the solution of a new crystal structure. The refinement of structure factors using molecular/dynamics, in programs like X-PLOR, works best with a reasonably accurate model as a starting point.[2] Therefore, it would be useful to have a method that could deduce the full coordinates from the α-carbon coordinates as closely as possible.

A method has been described for building a protein structure from its α-carbon coordinates.[3] The method is heavily based upon the use of existing structures as templates for the placement of the additional atoms, and a large amount of expertise in protein structure is required at each step in order to refine the structure. In a test case, flavodoxin (138 residues) was built and a backbone RMSD of 0.6 Å and an overall RMSD of 1.7 Å was attained.

It would be very useful to have a method that did not require this expert input that would be accessible to a larger number of investigators. In order to meet this need, a method that requires minimal expertise in protein structure on the part of the user has been developed. The procedure is based on the use of an automated molecular modeling procedure to build the coordinates of the unknown atoms. The model building is accomplished in two stages. With the α-carbon atom coordinates harmonically constrained, the backbone is built sequentially using only alanines, glycines, and prolines. The resulting partial structure is then refined using molecular dynamics. In the second stage the side chains are built by sequential addition and interim refining with molecular dynamics. This process results in good quality structures. For α lytic protease (2ALP, 198 residues)[3] a backbone RMSD deviation of 0.19 Å and an overall RMSD deviation of 1.24 Å are attained.

METHODS

All computations were performed with the CHARMm molecular modeling package from Polygen Corp.[5] and the GEMM program of Dr. Bernard Brooks.[6] The version 19 topology and parameter sets were used.[7] The distance-dependent dielectric option and a nonbonded cutoff of 9 Å were used. Calculations were performed on a Digital Equipment Corporation VAX 785 computer with an attached Star Technologies ST-50 array processor. Visual comparisons were performed on an Evans & Sutherland PS-390 using the MOGLI molecular graphics program. Sample input files are included in Appendix A.

The coordinates for α lytic protease (2ALP) and troponin C (2TNC)[8] were obtained from the Brookhaven Protein Data Bank.[1] The non-α-carbon coor-

Received June 14, 1989; revision accepted February 27, 1990.

Address reprint requests to Dr. Paul E. Correa, The Procter & Gamble Company, Miami Valley Laboratories, P.O. Box 398707, Cincinnati, Ohio 45239-8707.

TABLE I. Effects of Temperature and Time on Dynamics Refined 2ALP Backbone

	RMSD of backbone	
Temperature (°K)	50 psec	100 psec
500	.497	—
1000	.303	.272
1200	.353	—

dinates were removed from 2ALP. RMSD comparisons were made to the heavy atoms in the 2ALP entry without modification. The predicted coordinates for troponin C will be submitted to the Protein Data Bank.

RESULTS AND DISCUSSION

The first stage of the construction involves the building of the backbone. This is based on the idea that the information necessary for this process is contained in the α-carbon coordinates and the sequence. For this purpose, all amino acids except glycine and proline are changed to alanine. The prolines and glycines are required in order to represent the points of greater and lesser flexibility in the backbone. The reduction of the other amino acids to alanines was done to try to retain the information contained within the β-carbon positions, which serve to push the backbone gently in the right direction.

The crude backbone is built by sequentially adding residues restrained at their respective α-carbons. The first residue can be placed in a random orientation. Minimization of the structure after addition of the second residue orients it correctly. The α-carbon is harmonically restrained with a CHARMm force constant of 10 (20 kcal/mol) to remain close to its crystallographic position. Smaller restraints may be required if distortions of the amide linkages occur. The amide linkage for each residue is constrained to its *cis* or *trans* orientation as appropriate. The structure is then minimized using the steepest descents minimizer. This process is repeated until the whole backbone is built. It can be speeded up by starting at several points and building several chains simultaneously. This avoids the excessive repetitive minimization of the parts of the structure already built. The residues following glycine were chosen for starting points to minimize the effect of these flexible residues on the construction.

The resulting backbone is refined using molecular dynamics, retaining the α-carbon and amide linkage constraints. The backbone is heated to 1000°K for 100 psec of dynamics, which were the best conditions tried for 2ALP. Table I shows the effect of temperature and time on the RMSD fits of the backbone on the crystal structure. Optimal conditions may vary somewhat from molecule to molecule, and cannot be known in advance. However, variations in backbone RMSD are not very sensitive to these conditions (Tables I and V). At the end of the dynamics run the structure is cooled over 6 psec to 0°K and minimized using the CHARMm steepest descents algorithm. The cooling allows for atoms that have rotated to higher energy positions to return over modest energy barriers that would not be crossed with a simple minimization.

The effects of time and temperature can be interpreted in terms of the refinement purpose of the dynamics. Since the α-carbons are restrained, the temperature must be high in order to get a reasonable rate of conformational transitions. If the temperature is too high, then higher energy states are populated that cannot be escaped from upon cooling and minimization. The time is proportional to the level of equilibration attained. This process is being driven by the weak hydrogen bonding forces in the backbone. These will take time to work themselves into a consistent pattern. It is this establishment of an optimized hydrogen bonding network that is probably most responsible for the success of this method. In 2ALP most of the β-sheet interactions are formed during the dynamics refinement.

Cis-proline linkages can be built incorrectly and must be checked. The position of the *cis* linkages is determined by the distance between the β-carbons. If this is significantly less than 3.5 Å the linkage is assigned as *cis*. In 2ALP the *cis*-proline was the only out-of-phase residue after dynamics. If the proline and the preceding residue are deleted from the structure and rebuilt, they will be rebuilt correctly. Testing with other proteins shows that this procedure increases the accuracy of the original building procedure.

The construction of the side chains is more complicated than that for the backbone. The β-carbons have already been placed, as well as the proline rings. The side chains are built up sequentially by adding atoms at the next level out. First the γ level carbons, oxygens, and sulfurs (with their appended hydrogens) are added. The backbone is fixed and molecular dynamics at 800°K are run for 30 psec. Electrostatic charges on the nonbackbone atoms are scaled down with increasing distance from the geometric mean of the α-carbon coordinates. Charges in the 8–12 Å range from the geometric mean are scaled by 0.7; those over 12 Å are scaled by 0.3. This is used as a crude approximation of the solvation of the residues at the surface of the protein. The δ and ϵ atoms are then added in a similar manner. The balance of the atoms are added in the final step and the dynamics run at 1000°K for 30 psec. The all-atom model gives slightly better results for this part of the building procedure. Details of the coordinate manipulation commands can be found in the input files in the appendix. These can be adapted for use in other molecular mechanics programs. Other methods of side chain construction were tried, but these

TABLE II. RMS Fits to 2ALP of Built Structures*,†

	Crude backbone‡	After dynamics	Rebuild *cis*-proline
Backbone 1–197	0.823	0.272	0.182
Backbone and β-carbons	0.785	0.277	0.186
α-carbons	0.041	0.023	0.019
Nitrogens	0.382	0.170	0.085
Carbonyl carbons	0.501	0.216	0.148
β-carbons	0.568	0.297	0.201
Carbonyl oxygens	1.512	0.475	0.333

RMS Matches to 2ALP From Model Built With All Hydrogens Explicitly Present§

	Backbone	γ	δ	ϵ	All	Min SD	Reoriented
Backbone 1–197	0.182	0.182	0.182	0.182	0.182	0.182	0.182
with β-carbons	0.186	0.217	0.217	0.206	0.189	0.187	0.187
α-carbons	0.019	0.019	0.019	0.019	0.019	0.019	0.019
Nitrogens	0.085	0.085	0.085	0.085	0.085	0.085	0.085
Carbonyl C	0.148	0.148	0.148	0.148	0.148	0.148	0.148
β-C	0.201	0.334	0.333	0.293	0.214	0.206	0.206
Carbonyl O	0.333	0.333	0.333	0.333	0.333	0.333	0.333
All but H 1–197	0.199	0.717	1.022	1.122	1.291	1.289	1.248
Ala	0.183	0.221	0.231	0.216	0.193	0.185	0.185
Arg	0.228	0.751	1.113	1.351	2.303	2.304	2.304
Asp	0.176	0.731	1.283	1.057	0.798	0.792	0.477
Asn	0.268	0.799	1.745	1.605	1.578	1.577	1.551
Cys	0.166	1.043	1.146	0.861	0.833	0.836	0.836
Glu	0.158	0.820	1.304	1.284	1.369	1.370	1.283
Gln	0.143	0.661	0.914	1.461	1.151	1.144	1.095
Gly	0.201	0.201	0.201	0.201	0.201	0.201	0.201
His	0.171	0.426	0.390	1.875	2.698	2.701	2.701
Ile	0.130	1.166	1.221	1.264	1.115	1.117	1.117
Leu	0.164	0.614	1.294	1.239	1.297	1.293	1.293
Lys	0.077	0.283	0.467	0.894	0.811	0.808	0.808
Met	0.090	0.787	1.645	1.777	1.554	1.554	1.554
Phe	0.205	0.647	1.350	1.993	1.829	1.830	1.533
Pro	0.436	0.279	0.188	0.371	0.267	0.265	0.265
Ser	0.169	0.757	0.726	0.673	0.736	0.736	0.736
Thr	0.193	0.867	0.766	0.922	0.853	0.851	0.851
Tyr	0.157	0.498	1.388	1.796	2.363	2.357	2.131
Trp	0.171	0.741	1.699	1.424	1.696	1.688	1.688
Val	0.174	0.869	0.902	0.585	0.811	0.810	0.810

*RMS fitting was performed using the algorithm within CHARMm.
†Hydrogens were not included in the RMS match.
‡Crude backbone is after initial construction using minimization.
§For the fits for the individual residue types, the α-carbons were matched, then the individual RMS was performed without rotation.

invariably gave poorer results. This included the simplest method of building the backbone using complete residues instead of reducing the residues to alanine.

The RMSD fits for 2ALP at the various stages of construction are tabulated in Table II. The improvement of the backbone during dynamics is quite dramatic. Many amides flip 180° during the process to resolve conflicts that the original minimization process could not handle. The correction of the *cis*-proline brings the RMSD down to the final value of 0.19 Å. The carbonyl oxygens, which are the furthest from the axis between the α-carbons, have the poorest RMS fits of the backbone atoms, since it takes only a minor shift in the backbone angles to move these atoms considerably. The side chain RMSD values are given at the various stages of construction and are detailed by residue type as well. The greatest errors are found with the aromatic (Phe, Tyr, Trp, and His) and charges (Arg, Lys) residues. In the earlier study[3] it was also noticed that the aromatic and charged residues were the hardest to place. A stereo view of the overlap of the model built and the crystallographic coordinates is shown in Figure 1. The major difference is the placement of the histidine in another Chi 1 rotameric state. Protonation of the histidine could correct this, providing an electrostatic attraction to the aspartic acid. The other two members of the catalytic triad, Ser-195 and Asp-102, are accurately placed.

The quality of the fit decreases with increasing distance from the center. Table III shows the values for the RMSD fits based on the distance from the geometric mean of the α-carbon coordinates. This can be attributed to the greater number of stable conformations in the less restricted surface environment. The distribution of various residue types (surface vs. interior) affects the pattern for individual

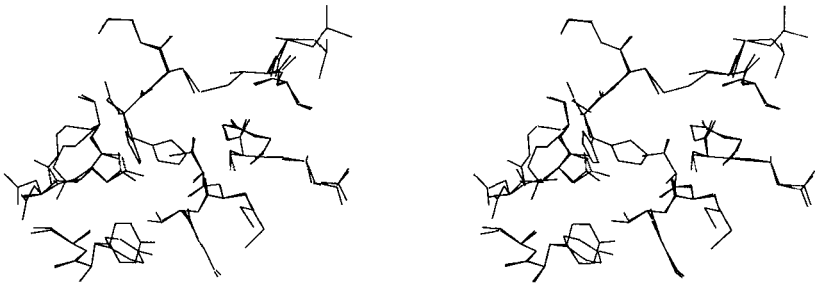

Fig. 1. Stereo view of the overlap of the model built and crystallographic coordinates of α lytic protease. Included are residues 41–43, 56–59, 101–103, 171–173, and 194–197.

residues. The number of atoms in each level increases dramatically, as would be expected.

The difference between the RMSD fits for the backbone and the overall model (a difference of 1.1 Å) is comparable to that reported in the earlier study.[3] This has some implications for further work in this area. It is possible that the more expert method employed cannot do much better than a dynamics sampling of the conformational space as used above. In other words, the positioning of the side chains may be much more intimately related to the tertiary structure than to the primary and secondary structure, which are the criteria for selecting fragments in the more expert methods. If this is the case, the use of comparable fragments may not be the way to go in predicting positions. The overall Chi 1 fits are reasonable. Out of 142 residues with rotatable side chains, 88 have Chi 1 within 60° of the correct value. Of the incorrect residues 36 were off by more than 60° with a negative rotation, 18 were off with a positive rotation. The RMSD of Chi 1 is 75°, with a mean error of −16°. The use of more spatially oriented data that take into account the three-dimensional packing of these structures may be more suitable. This remains an unresolved question at this point.

The construction of the 2TNC model was accomplished in a similar manner. Since there are no *cis* linkages the procedure was simplified in this respect. The RMSD fit is 0.41 Å for the main chain atoms (0.34 Å if three particularly poorly fitting atoms are eliminated). This number is not as good as that attained with 2ALP, but it must be remembered that this calculation was performed with unrefined α-carbon coordinates. The α-carbon coordinates in the set used for construction had an RMSD of 0.30 Å compared to the final refined version. Thus, it can be seen that the placement of the other main chain atoms is accurate with respect to the placement of the α-carbons. The side chains reflect this problem even more. The RMSD fit for all 1234 atoms is 1.68 Å. The major deviations that were observed involve the positioning of Phe-22 and Phe-75 (which rotated in concert with each other), Phe-151, and Phe-13. The charged residues were also poor in their fits, mainly due to excessive folding back onto the backbone. This is very much in accord with what we observe in the 2ALP structure. The high surface area for troponin C probably also contributes to the larger RMSD error since the method does not do as well in surface regions, where structural constraints are minimal.

In order to further test this method, the construction of flavodoxin from the α-carbons of 3FXN[9] was performed. The RMS deviations are summarized in Tables IV and V. The dynamics refined structure for the backbone is closer to the crystal structure than the one previously reported[3] (RMSD of 0.49 Å vs. 0.6 Å). This is not of the same quality as obtained in the attempts with 2ALP and 2TNC, although the effect of the cofactor, which was not included in the calculation, cannot be ruled out. The high quality of the backbone in helices and sheets is a reflection of the importance of the hydrogen bonding in determining the orientation of these residues. The RMSD for the overall structure is comparable to that obtained previously[3] (RMSD of 1.64 Å here vs. 1.73 Å). It is interesting to note that a better fit to the crystal structure is obtained for the β-sheet region[3] (RMSD of 1.08 Å here vs. 1.86 Å), which occupies the center of the molecule. The better fit probably results from the limitations in packing arrangements explored by the dynamics. At the surface, where there is often little or no constraint, many of the side chains are in different rotamer positions resulting in a substantially poorer fit to the crystal structure. The effects of crystal packing and solvent organization probably play a major role in determining the observed orientations, and of course are not reproduced here. Finally, any error in this smaller structure has a larger impact on the total RMSD fits.

The computer time required for these calculations is not inconsiderable. The use of the ST-50 makes the process fairly rapid, but certainly super- and minisupercomputers are in this range. The construction of the initial backbone structure by minimization takes less than an hour. The 100 psec of dynamics refinement takes about 12–15 hours (50 psec gives comparable accuracy in half the time). The construction of the side chains takes from 18 to 30

TABLE III. RMS Fits for Final 2ALP Coordinates: α-Carbons Matched, Then Comparisons Made in Place*

	Distance from center†			
	0–4 Å	4–8 Å	8–12 Å	>12 Å
Before orienting	0.082	1.020	1.157	1.365
After orienting	0.092	0.872	1.057	1.343
Ala	0.000	0.095	0.177	0.189
Arg	0.000	1.768	0.923	2.624
Asp	0.000	0.652	0.371	0.000
Asn	0.000	0.000	0.000	1.551
Cys	0.000	0.364	0.270	1.093
Glu	0.000	0.088	0.804	1.533
Gln	0.000	0.000	0.189	1.167
Gly	0.078	0.083	0.307	0.000
His	0.000	0.000	0.786	4.161
Ile	0.000	0.239	1.148	1.113
Leu	0.000	0.000	1.190	1.332
Lys	0.000	0.000	0.000	0.808
Met	0.038	1.184	1.447	2.091
Phe	0.000	1.605	2.194	0.604
Pro	0.000	0.000	0.000	0.265
Ser	0.055	0.563	0.900	0.746
Thr	0.000	0.163	0.141	0.886
Tyr	0.000	0.233	0.444	2.482
Trp	0.000	0.347	1.494	2.652
Val	0.103	0.992	1.124	0.693

*RMS fits were performed using the algorithm within CHARMm.
†Distances were calculated from the geometric mean of the α-carbon coordinates.

TABLE IV. RMS Fits During the Construction of 3FXN From α-Carbon Coordinates

	Crude backbone*	Refined backbone†	γ	δ	ϵ	All
Backbone 1–137	0.67	0.49	0.49	0.49	0.49	0.49
Backbone and β-carbons 1–137	0.64	0.47	0.48	0.48	0.48	0.47
α-Carbons	0.04	0.03	0.03	0.03	0.03	0.03
Nitrogens	0.27	0.20	0.20	0.20	0.20	0.20
Carbonyl carbons	0.44	0.30	0.30	0.30	0.30	0.30
β-Carbons	0.53	0.33	0.40	0.44	0.43	0.35
Carbonyl O res 1–137	1.23	0.91	0.91	0.91	0.91	0.91
All but hydrogens 1–137	0.64	0.47	0.81	1.14	1.49	1.64
Ala	0.55	0.14	0.14	0.16	0.21	0.13
Arg	0.09	0.07	0.97	0.99	1.41	3.51
Asp	0.77	0.63	0.86	1.53	1.48	1.64
Asn	0.75	0.65	0.91	1.48	1.69	1.59
Cys	0.28	0.22	0.95	0.98	0.35	0.23
Glu	0.64	0.53	0.90	1.31	2.03	2.08
Gln	0.18	0.10	1.10	1.33	1.67	2.01
Gly	0.61	0.69	0.69	0.69	0.69	0.69
Ile	0.84	0.14	0.64	0.98	1.16	1.06
Leu	0.28	0.27	0.52	1.38	1.30	1.03
Lys	0.31	0.33	0.72	0.92	1.36	2.07
Met	0.53	0.27	0.59	0.88	1.49	1.31
Phe	0.21	0.14	0.64	1.21	1.26	0.94
Pro	0.99	0.74	0.69	0.72	0.69	0.71
Ser	0.68	0.44	0.88	0.89	0.93	0.94
Thr	0.86	0.83	1.24	1.12	1.29	1.11
Tyr	0.41	0.15	0.61	0.91	1.38	0.93
Trp	0.80	0.77	1.13	1.46	3.05	3.72
Val	0.63	0.23	0.89	0.98	1.01	0.74

*Crude backbone is after initial construction using minimization.
†Refined backbone is after dynamics refinement.

TABLE V. RMS Fits During Construction of 3FXN From α-Carbon Coordinates

	Helices	β-Sheets
Crude backbone	0.44	0.47
Refined backbone	0.27	0.17
γ	0.68	0.72
δ	0.93	1.21
ϵ	1.39	1.12
All	1.63	1.08

hours. The side chain construction could be cut significantly if the united atom model is used, with a minimal loss in accuracy (0.1 in the overall RMSD). Our experience with X-PLOR indicates that the initial dynamics refinement of the backbone is useful since X-PLOR tends to have difficulties with out of phase amide linkages. Rotations of the backbone tend to be particularly difficult.[2] X-PLOR does refine the side chains more efficiently, probably because these are rotationally more mobile. The methodology described here could be incorporated in a program such as X-PLOR to build the backbone first without side chains to allow generation of a more accurate backbone. Subsequent side chain construction and refinement using X-ray data may circumvent some of the problems described in the annealing process.[2]

Several extensions of the method remain untested. The use of published stereo views facilitates the accurate placement of side chains in the absence of side chain coordinates. This is made possible by the accuracy of the backbone built from the α-carbon coordinates. Adding the side chains usually has only one possibility at this point. It may also be useful to include the small hydrogen bonding side chains (Ser, Asn, Asp) during backbone construction to give a more complete hydrogen bonding network without impairing backbone flexibility. If disulfide bridges are known, or suspected, they could be included as well. The harmonic constraint force could also be made inversely proportional to the temperature factors. The details of variations of this methodology and its extension will be the subject of future reports.

In conclusion, a method has been developed for the accurate generation of protein backbones from their corresponding α-carbon coordinates. This method involves sequential residue generation and minimization. The final structure is refined using constained molecular dynamics. Side chains can be added using a molecular dynamics stepwise approach to give a reasonable approximation to their location and conformation. Further refinement by inspection of the final model by an expert in protein structure is possible. The importance of this method lies in its lack of protein structure expert input to arrive at a good approximation of the protein structure. This should be of considerable usefulness for many areas of research, especially drug design. In addition this procedure will be of use for protein structure determinations. If the map is good enough to establish α-carbon coordinates then the crystallographer can generate an approximate structure using this methodology. The initial model thus generated can be subjected to structure factor refinement using molecular dynamics. This type of refinement can be extremely useful for the rapid refinement of a structure.[10]

ACKNOWLEDGMENTS

We are grateful to Dr. Michael James and Natalie Strynadka for the RMSD fits to the troponin C structure before its publication and general discussions on these structures, and to Dr. Bernard Brooks for discussions on the use of molecular dynamics.

REFERENCES

1. Bernstein, F. C., Koetzle, T. F., Williams, E. J. B., Meyer, E. F., Jr., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A Computer- based archival file for molecular structures. J. Mol. Biol. 112:535–542, 1977.
2. Kuriyan, J., Brunger, A. T., Karplus, M., Hendrickson, W. A. X-Ray refinement of protein structures by simulated annealing: Test of the method on myohemerythrin. Acta Crys. A45:396–409, 1989.
3. Reid, L. S., Thornton, J. M. Prediction of protein structure from Cα atomic coordinates. In: "Protein Structure, Folding, and Design 2." New York: Liss, 1987: 93–102.
4. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D., James, M. N. G. Refined structure of alpha-lytic protease at 1.7 angstroms resolution. Analysis of hydrogen bonding and solvent structure. J. Mol. Biol. 184:479–502, 1985.
5. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M. CHARMM: A program for macaromolecular energy, minimization and dynamics calculations. J. Comput. Chem. 4:187–217, 1983.
6. Brooks, B. R. Personal communication, 1987.
7. Reiher, W. E. Dissertation, Harvard University, Cambridge, MA, 1984.
8. Herzberg, O., James, M. N. G. Structure of the calcium regulatory muscle protein troponin-C at 2.8 angstroms resolution. Nature (London) 313:653–659, 1985.
9. Smith, W. W., Burnett, R. M., Darling, G. D., Ludwig, M. L. Structure of the semiquinone form of flavodoxin from clostridium MP. Extension 1.8 angstroms resolution and some comparisons with the oxidized state. J. Mol. Biol. 117:195–225, 1977.
10. James, M. N. G. Personal communication, 1988.

APPENDIX. SAMPLE INPUT FILES

Part 1. Building the Crude Backbone

```
* BUILDING 2TNC BACKBONE ON THE STAR
*

READ RTF UNIT 1
READ PARAMETER UNIT 2

OPEN UNIT 14 NAME 2TNCSTARTFAST.PDB
READ CARD

! SET PARAMETERS TO GLYCINE POSITIONS
! TO INITIATE THE BUILDING OF MULTIPLE
! CHAINS ALONG THE BACKBONE
SET 0 6
SET 1 37
```

```
SET 2 45
SET 3 52
SET 4 73
SET 5 94
SET 6 113
SET 7 121
SET 8 128
SET 9 149

LABEL LOOP

!READ IN PRECONSTRUCTED PSF AND
!CONSTRAINT FILES SO THESE WILL NOT
!NEED TO BE
!RECONSTRUCTED AT EVERY LOOP CYCLE
OPEN UNIT 3 NAME 2TNC.PSF READ UNFORM
OPEN UNIT 10 NAME 2TNCCONS.CON READ -
UNFORM
OPEN UNIT 13 NAME 2TNCCA.PDB READ -
CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
OPEN UNIT 4 NAME 2TNCIC.IC READ CARD
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10

!DELETE ATOMS FROM PSF THAT ARE NOT TO
!BE BUILT THIS CYCLE
IF 0 GE 35 GOTO ENDLOOP
DELE ATOM SELE ((.NOT. (RESID 5:@0)) .AND. -
(RESID 5:35))END
IF 1 GE 43 GOTO TWO
DELE ATOM SELE ((.NOT. (RESID 36@1)) .AND. -
(RESID 36:43))END
LABEL TWO
IF 2 GE 50 GOTO THREE
DELE ATOM SELE ((.NOT. (RESID 44:@2)) .AND. -
(RESID 44:50))END
LABEL THREE
IF 3 GE 71 GOTO FOUR
DELE ATOM SELE ((.NOT. (RESID 51:@3)) .AND. -
(RESID 51:71))END
LABEL FOUR
IF 4 GE 92 GOTO FIVE
DELE ATOM SELE ((.NOT. (RESID 72:@4)) .AND. -
(RESID 72:92))END
LABEL FIVE
IF 5 GE 111 GOTO SIX
DELE ATOM SELE ((.NOT (RESID 93:@5)) .AND. -
(RESID 93:111))END
LABEL SIX
IF 6 GE 119 GOTO SEVEN
DELE ATOM SELE ((.NOT. (RESID 112:@6)) -
.AND. (RESID 112:119))END
LABEL SEVEN
IF 7 GE 126 GOTO EIGHT
DELE ATOM SELE ((.NOT. (RESID 120:@7)).
.AND. (RESID 120:126))END
LABEL EIGHT
IF 8 GE 147 GOTO NINE
DELE ATOM SELE ((.NOT. (RESID 127:@8)) -
.AND. (RESID 127:147))END
LABEL NINE
IF 9 GE 161 GOTO TEN
DELE ATOM SELE ((.NOT. (RESID 148:@9)) -
.AND. (RESID 148:161))END
LABEL TEN

!READ IN PLACED OR PREVIOUSLY BUILT
!COORDINATES
READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14

!CONSTRUCT UNKNOWN COORDINATES
IC FILL PRESERVE
IC PARAM
IC BUILD
NBON RDIEL

!CONSTRAIN ALPHA CARBON POSITIONS
CONS HARM SELE ATOM * * CA END FORCE 10 -
COMP MASS

!INVOKE STAR PROCESSOR AND MINIMIZE
STAR INBF 20 DOUB -
     ATOM RDIE SHIF VSHI VDIS CUTNB 9.0
MINI SD NPRINT 50 NSTEP 300
STAR OFF

OPEN WRITe CARD UNIT 14 NAME
2TNCBUILDFAST.PDB
WRITe COOR PDB UNIT 14
* 2TNC PARTIAL STRUCTURE CHARMM -
BUILT MIN 300 SD OMEGA CONSTRAINED
*

CLOSE UNIT 14
!CLEAR CONSTRAINTS AND PSF FOR NEXT
!CYCLE
CONS HARM SELE ALL END FORCE 0.0
DELE ATOM SELE ALL END
CONS CLDH
OPEN READ CARD UNIT 14 NAME -
2TNCBUILDFAST.PDB

!ADJUST COUNTERS FOR NEXT CYCLE
INCREMENT 0 BY 1
INCREMENT 1 BY 1
INCREMENT 2 BY 1
INCREMENT 3 BY 1
INCREMENT 4 BY 1
INCREMENT 5 BY 1
INCREMENT 6 BY 1
```

```
INCREMENT 7 BY 1
INCREMENT 8 BY 1
INCREMENT 9 BY 1
GOTO LOOP
LABEL ENDLOOP

! DO FINAL MINIMIZATION
OPEN UNIT 3 NAME 2TNC.PSF READ UNFORM
OPEN UNIT 10 NAME 2TNCCONS.CON READ -
UNFORM
OPEN UNIT 13 NAME 2TNCCA.PDB READ -
CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
OPEN UNIT 4 NAME 2TNCIC.IC READ CARD
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10
READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14

IC FILL PRESERVE
IC PARAM
IC BUILD
NBON RDIEL

CONS HARM SELE ATOM * * CA END FORCE -
10 COMP MASS

STAR INBF 20 DOUB -
  ATOM RDIE SHIF VSHI VDIS CUTNB 9.0
MINI SD NPRINT 50 NSTEP 300
STAR OFF

OPEN WRITe CARD UNIT 14 NAME
2TNCBUILDFAST.PDB
WRITe COOR PDB UNIT 14
* 2TNC PARTIAL STRUCTURE CHARMM -
BUILT MIN 300 SD OMEGA CONSTRAINED
*

CLOSE UNIT 14
CONS HARM SELE ALL END FORCE 0.0
CONS CLDH
ENERGY

STOP
```

Part 2. Refinement of the Backbone Through Dynamics

```
* BUILDING COMPLETE 2TNC ON THE STAR -
DYNAMICS
*
OPEN UNIT 1 NAME SYS$US3:[BROOKS. -
CHARMM.TOPPAR]TOPH19.MOD READ -
UNFORM
READ RTF FILE UNIT 1
CLOSE UNIT 1

OPEN UNIT 2 NAME SYS$US3:[BROOKS. -
CHARMM.TOPPAR]PARAM19.MOD READ -
UNFORM
READ PARAMETER FILE UNIT 2
CLOSE UNIT 2

OPEN UNIT 14 NAME 2TNCBUILDFAST.PDB
READ CARD

!READ IN PRECONSTRUCTED PSF AND
!CONSTRAINT FILES
OPEN UNIT 3 NAME 2TNC.PSF READ UNFORM
OPEN UNIT 10 NAME 2TNCCONS.CON READ
UNFORM
OPEN UNIT 13 NAME 2TNCCA.PDB READ
CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
OPEN UNIT 4 NAME 2TNCIC.IC READ CARD
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10
IC PURGE

!READ IN CRUDE CONSTRUCTED BACKBONE
!AND CRYSTALLOGRAPHIC ALPHA CARBON
!POSITIONS
READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14

NBON RDIEL

CONS HARM SELE ATOM * * CA END FORCE -
10 COMP MASS

! INVOKE STAR PROCESSOR AND RUN
!DYNAMICS
STAR INBF 20 DOUBLE DPI0 ECHECK 1.0E4
!MINI SD NSTEP 50 NPRINT 10
OPEN WRITE UNIT 50 FILE NAME
2TNCFASTDYNHEAT1000-2.TRJ
DYNA VERLET STRT NSTEP 4000 TIMESTEP -
→0.001 - IPRFRQ 1000 IHTFRQ 100 IEQFRQ 0 -
  NTRFRQ 1000 -
  IUNCRD 50 AVERAGE ISEED 314159 -
  NPRINT 1000 NSAVC 1000 NSAVV 0 INBFRQ -
  25 IHBFRQ 0 -
  FIRSTT 0.0 FINALT 1000.0 TEMINC 30.0 -
```

```
  IASORS 1 IASVEL 1 ISCVEL 0 ICHECW 0 -
  TWINDH 10.0 TWINDL -10.0


OPEN WRITE UNIT 50 FILE NAME
2TNCFASTDYN1000-2.TRJ
DYNA VERLET STRT NSTEP 100000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 0 IEQFRQ 0 -
  NTRFRQ 1000 - IUNCRD 50 AVERAGE -
  ISEED 314159 - NPRINT 1000 NSAVC 1000 -
  NSAVV 0 INBFRQ 25 IHBFRQ 0 - FIRSTT -
  1000.0 FINALT 1000.0 TEMINC 0.0 - IASORS 0 -
  IASVEL 1 ISCVEL 0 ICHECW 0 TWINDH -
  10.0 TWINDL -10.0


OPEN WRITE UNIT 32 CARD NAME -
2TNCFASTDYN1000-2.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 1000K DYN 100 PS INITIAL
*
OPEN WRITE UNIT 32 CARD NAME -
2TNCFASTDYN1000-2.VEL2
WRITE COOR CARD COMP UNIT 32
* FINAL VELOCITY FOR 2TNC STAR VERSION
*


OPEN WRITE UNIT 50 FILE NAME
2TNCFASTDYN0-2.TRJ
DYNA VERLET STRT NSTEP 3300 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 100 IEQFRQ 0 -
  NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
  314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
  INBFRQ 25 IHBFRQ 0 - FIRSTT 1000.0 FINALT -
  0.0 TEMINC -30.0 - IASORS 1 IASVEL 1 ISCVEL -
  0 ICHECW 0 TWINDH 10.0 TWINDL -10.0


OPEN WRITE UNIT 32 CARD NAME -
2TNCFASTDYN0-2.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION OK DYN COOLED
*
OPEN WRITE UNIT 32 CARD NAME -
2TNCFASTDYNO-2.VEL2
WRITE COOR CARD COMP UNIT 32
* FINAL VELOCITY FOR 2TNC STAR VERSION
*

!MINIMIZE FINAL STRUCTURE FOR
!BACKBONE
MINI SD NPRINT 50 NSTEP 500
OPEN WRITE UNIT 32 CARD NAME -
2TNCFASTDYN1000-2MINSD.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 1000K DYN 100 PS FINAL MIN 500 SD
*

STOP
```

Part 3. Building the Side Chains

```
* BUILDING COMPLETE 2TNC ON THE STAR -
DYNAMICS
*

READ RTF UNIT 1
READ PARAMETER UNIT 2

OPEN UNIT 14 NAME 2TNCBACKBONE.PDB
READ CARD

!READ IN PRECONSTRUCTED ALL ATOM PSF
!AND CONSTRAINT FILES
OPEN UNIT 3 NAME 2TNCH.PSF READ -
UNFORM
OPEN UNIT 10 NAME 2TNCCONSH.CON READ -
UNFORM
OPEN UNIT 13 NAME SYS$US3:[CORREA.
TNC]2TNCCAO.PDB READ CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
OPEN UNIT 4 NAME 2TNCICH.IC READ CARD
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10

!READ IN COORDINATES
READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14

!BUILD ALL ATOMS IN THE SIDE CHAINS
IC FILL PRESERVE
IC PARAM
IC BUILD

!DELETE ALL ATOMS BEYOND THE FIRST
!LEVEL CONSTRUCTED
DELE ATOM SELE (TYPE CD* .OR. TYPE OD* -
.OR.- TYPE ND* .OR. TYPE HD* .OR. -
  TYPE SD .OR. TYPE CE* .OR. TYPE OE* .OR. -
  TYPE NE* .OR. TYPE HE* .OR. -
  TYPE CZ* .OR. TYPE NZ .OR. TYPE HZ* .OR. -
  TYPE NH* .OR. TYPE HH* .OR. -
  TYPE OH .OR. TYPE CH2) .AND. (.NOT. -
  (RESNAME PRO)) END

!SCALE CHARGES TO SIMULATE SOLVATION
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 8.0) .AND. - (POINT 23.76 27.57 -
```

```
   14.68 CUT 12.0)) .AND. (.NOT. (TYPE - H .OR. -
   TYPE O)) END MULT 0.7
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 12.0) .AND. - (POINT 23.76 -
   27.57 14.68 CUT 30.0)) .AND. (.NOT. (TYPE - H -
   .OR. TYPE O)) END MULT 0.3

NBON RDIEL
!FIX BACKBONE
CONS FIX SELE (TYPE N .OR. TYPE CA .OR.
TYPE C .OR. TYPE O) END

!INVOKE STAR PROCESSOR AND RUN
!DYNAMICS
STAR INBF 20 DOUBLE DPI0 ECHECK 1.0E3
MINI SD NPRINT 10 NSTEP 60 !INITIAL
!MINIMIZATION TO RELIEVE BAD CONTACTS
OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDEHEAT1.TRJ
DYNA VERLET STRT NSTEP 3000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 100 IEQFRQ 0 -
   NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
   314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
   INBFRQ 25 IHBFRQ 0 -FIRSTT 0.0 FINALT -
   800.0 TEMINC 30.0 - IASORS 1 IASVEL 1 -
   ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
   -10.0

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-800-1.TRJ
DYNA VERLET STRT NSTEP 30000 TIMESTEP -
0.001 - IPRFRQ 5000 IHTFRQ 0 IEQFRQ 0 -
   NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
   314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
   INBFRQ 25 IHBFRQ 0 - FIRSTT 800.0 FINALT -
   800.0 TEMINC 0.0 - IASORS 0 IASVEL 1 -
   ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
   -10.0

OPEN WRITE UNIT 32 CARD NAME 2TNCSIDE- -
800-1.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 800K DYN 30 PS GAMMA ATOMS -
BUILT
*
OPEN WRITE UNIT 32 CARD NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-800-
   1.VEL2 -
WRITE COOR CARD COMP UNIT 32
* FINAL VELOCITY FOR 2TNC STAR VERSION
*
! CLEAR CONSTRAINTS AND PSF FOR NEXT
!CYCLE OF BUILDING SIDE CHAINS
CONS FIX SELE NONE END
DELE ATOM SELE ALL END

!REPEAT THE PROCESS ABOVE UNTIL ALL
!ATOMS ARE BUILT
OPEN UNIT 3 NAME 2TNCH.PSF READ -
UNFORM
OPEN UNIT 10 NAME 2TNCCONSH.CON READ -
UNFORM
OPEN UNIT 13 NAME SYS$US3:[CORREA. -
TNC]2TNCCAO.PDB READ CARD
OPEN UNIT 4 NAME 2TNCICH.IC READ CARD
OPEN UNIT 14 NAME 2TNCSIDE-800-1.PDB
READ CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10
IC PURGE

READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14
IC PURGE
IC FILL PRESERVE
IC PARAM
IC BUILD

DELE ATOM SELE ( TYPE CE* .OR. TYPE OE* -
→.OR. TYPE NE* .OR. TYPE HE* .OR. - TYPE CZ* -
   .OR. TYPE NZ .OR. TYPE HZ* .OR. -
   TYPE NH* .OR. TYPE HH* .OR. -
   TYPE OH .OR. TYPE CH2 .OR. TYPE HD1) END
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 8.0) .AND. - (POINT 23.76 -
   27.57 14.68 CUT 12.0)) .AND. (.NOT. (TYPE -
   (H .OR. TYPE O)) END MULT 0.7
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 12.0) .AND. - (POINT 23.76 -
   27.57 14.68 CUT 30.0)) .AND. (.NOT. (TYPE -
   H. .OR. TYPE O)) END MULT 0.3

NBON RDIEL

CONS FIX SELE (TYPE N .OR. TYPE CA .OR. -
TYPE C .OR. TYPE O) END

STAR INBF 20 DOUBLE DPI0 ECHECK 1.0E3
MINI SD NPRINT 10 NSTEP 60

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDEHEAT2.TRJ
DYNA VERLET STRT NSTEP 3000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 100 IEQFRQ 0 -
   NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
   314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
   INBFRQ 25 IHBFRQ 0 - FIRSTT 0.0 FINALT -
   800.0 TEMINC 30.0 - IASORS 1 IASVEL 1 -
```

```
   ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
   -10.0

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-800-2.TRJ
DYNA VERLET STRT NSTEP 30000 TIMESTEP -
0.001 - IPRFRQ 5000 IHTFRQ 0 IEQFRQ 0 -
   NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
   314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
   INBFRQ 25 IHBFRQ 0 - FIRSTT 800.0 FINALT -
   800.0 TEMINC 30.0 - IASORS 0 IASVEL 1 -
   ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
   -10.0

OPEN WRITE UNIT 32 CARD NAME 2TNCSIDE -
800-2.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 800K DYN 30 PS DELTA ATOMS -
BUILT
*
OPEN WRITE UNIT 32 CARD NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-800-2.VEL2
WRITE COOR CARD COMP UNIT 32
* FINAL VELOCITY FOR 2TNC STAR VERSION
*
CONS FIX SELE NONE END
DELE ATOM SELE ALL END

OPEN UNIT 3 NAME 2TNCH.PSF READ -
UNFORM
OPEN UNIT 10 NAME 2TNCCONSH.CON READ -
UNFORM
OPEN UNIT 13 NAME SYS$US3:[CORREA.TNC] -
2TNCCAO.PDB READ CARD
OPEN UNIT 4 NAME 2TNCICH.IC READ CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10
IC PURGE
OPEN UNIT 14 NAME 2TNCSIDE-800-2.PDB
READ CARD READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14
IC PURGE
IC FILL PRESERVE
IC PARAM
IC BUILD

DELE ATOM SELE (TYPE CZ* .OR. TYPE NZ .OR. -
→TYPE HZ* .OR. TYPE NH* .OR. TYPE HH* -
   .OR. TYPE OH .OR. TYPE CH2 .OR. TYPE HE) -
   END

SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 8.0) .AND. - (POINT 23.76 -
   27.57 14.68 CUT 12.0)) .AND. -
   (.NOT. (TYPE -
   H .OR. TYPE O)) END MULT 0.75
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 12.0) .AND. - (POINT 23.76 -
   27.57 14.68 CUT 30.0)) .AND. -
   (.NOT. (TYPE -
   H. .OR. TYPE O)) END MULT 0.5

NBON RDIEL

CONS FIX SELE (TYPE N .OR. TYPE CA .OR. -
TYPE C .OR. TYPE O) END

STAR INBF 20 DOUBLE DPI0 ECHECK 1.0E3
MINI SD NPRINT 10 NSTEP 60

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDEHEAT3.TRJ
DYNA VERLET STRT NSTEP 3000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 100 IEQFRQ 0 -
   NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
   314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
   INBFRQ 25 IHBFRQ 0 - FIRSTT 0.0 FINALT -
   800.0 TEMINC 30.0 - IASORS 1 IASVEL 1 -
   ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
   -10.0

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-800-3.TRJ
DYNA VERLET STRT NSTEP 30000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 0 IEQFRQ 0 -
   NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
   314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
   INBFRQ 25 IHBFRQ 0 - FIRSTT 800.0 FINALT -
   800.0 TEMINC 30.0 - IASORS 0 IASVEL 1 -
   ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
   -10.0

OPEN WRITE UNIT 32 CARD NAME 2TNCSIDE- -
800-3.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 800K DYN 30 PS EPSILON ATOMS -
BUILT
*
OPEN WRITE UNIT 32 CARD NAME SYS$SCRA:
-
[SCRATCH.CORREA.TNC]2TNCSIDE-800-3.VEL2
WRITE COOR CARD COMP UNIT 32
* FINAL VELOCITY FOR 2TNC STAR VERSION
*
CONS FIX SELE NONE END
DELE ATOM SELE ALL END

OPEN UNIT 3 NAME 2TNCH.PSF READ -
UNFORM
```

```
OPEN UNIT 10 NAME 2TNCCONSH.CON READ -
UNFORM
OPEN UNIT 13 NAME SYS$US3:[CORREA. -
TNC]2TNCCAO.PDB READ CARD
OPEN UNIT 4 NAME 2TNCICH.IC READ CARD
READ PSF FILE UNIT 3
CLOSE UNIT 3
IC READ UNIT 4
CLOSE UNIT 4
READ CONS FILE UNIT 10
CLOSE UNIT 10
IC PURGE
OPEN READ UNIT 14 NAME 2TNCSIDE -
800-3.PDB CARD

READ COOR PDB UNIT 13
COOR COPY COMP
CLOSE UNIT 13
READ COOR PDB UNIT 14
CLOSE UNIT 14
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 8.0) .AND. - (POINT 23.76 -
  27.57 14.68 CUT 12.0)) .AND. (.NOT. (TYPE - H -
  .OR. TYPE O)) END MULT 0.75
SCALAR CHARGE SELE ((.NOT. POINT 23.76 -
→27.57 14.68 CUT 12.0) .AND. - (POINT 23.76 -
  27.57 14.68 CUT 30.0)) .AND. -
  (.NOT. (TYPE -
  H .OR. TYPE O)) END MULT 0.5

IC PURGE
IC FILL PRESERVE
IC PARAM
IC BUILD

NBON RDIEL

CONS FIX SELE (TYPE N .OR. TYPE CA .OR. -
TYPE C .OR. TYPE O) END

STAR INBF 20 DOUBLE DIP0 ECHECK 1.0E3
MINI SD NPRINT 10 NSTEP 60

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDEHEAT4.TRJ
DYNA VERLET STRT NSTEP 8000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 100 IEQFRQ 0 -
  NTRFRQ 1000 -
  IUNCRD 50 AVERAGE ISEED 314159 -
  NPRINT 50 NSAVC 1000 NSAVV 0 INBFRQ 25-
  IHBFRQ 0 -
  FIRSTT 0.0 FINALT 1000.0 TEMINC 30.0 -
  IASORS 1 IASVEL 1 ISCVEL 0 ICHECW 0 -
  TWINDH 10.0 TWINDL -10.0

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-1000-4.TRJ
DYNA VERLET STRT NSTEP 30000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 0 IEQFRQ 0 -
  NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
  314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
  INBFRQ 25 IHBFRQ 0 - FIRSTT 1000.0 -
  FINALT 1000.0 TEMINC 30.0 - IASORS 0 -
  IASVEL 1 ISCVEL 0 ICHECW 0 TWINDH 10.0 -
  TWINDL -10.0

OPEN WRITE UNIT 50 FILE NAME SYS$SCRA: -
[SCRATCH.CORREA.TNC]2TNCSIDE-COOL.TRJ
DYNA VERLET STRT NSTEP 8000 TIMESTEP -
→0.001 - IPRFRQ 5000 IHTFRQ 100 IEQFRQ 0 -
  NTRFRQ 1000 - IUNCRD 50 AVERAGE ISEED -
  314159 - NPRINT 50 NSAVC 1000 NSAVV 0 -
  INBFRQ 25 IHBFRQ 0 - FIRSTT 1000.0 FINALT -
  0.0 TEMINC -30.0 - IASORS 0 IASVEL 1 -
  ISCVEL 0 ICHECW 0 TWINDH 10.0 TWINDL -
  -10.0

OPEN WRITE UNIT 32 CARD NAME 2TNCSIDE-
1000-4.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 1000K DYN 30 PS ZETA ATOMS -
BUILT
*
OPEN WRITE UNIT 32 CARD NAME SYS$SCRA:
[SCRATCH.CORREA.TNC]2TNCSIDE-1000-4. -
VEL2
WRITE COOR CARD COMP UNIT 32
* FINAL VELOCITY FOR 2TNC STAR VERSION
*
CONS FIX SELE NONE END
ENERGY
MINI SD NSTEP 500 NPRINT 50
ENERGY
OPEN WRITE UNIT 32 CARD NAME 2TNCSIDE- -
1000-4MINSD.PDB
WRITE COOR PDB UNIT 32
* FINAL COORDINATES FOR 2TNC STAR -
VERSION 1000K DYN 30 PS BUILT MIN SD
*

STOP
```